

EMANUELLY CANABRAVA MAGALHÃES

**MODELAGEM DA PRODUÇÃO DE *Elaeis guineensis* Jacq. UTILIZANDO REDES
NEURONAIS ARTIFICIAIS**

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

Orientador: Helio Garcia Leite

Coorientadores: Carlos Alberto Araújo Júnior
Carlos Pedro Boechat Soares

**VIÇOSA – MINAS GERAIS
2022**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

M188m
2022
Magalhães, Emanuely Canabrava, 1996-
Modelagem da produção de *Elaeis guineensis* Jacq.
utilizando redes neuronais artificiais / Emanuely Canabrava
Magalhães. – Viçosa, MG, 2022.
1 dissertação eletrônica (76 f.): il.

Orientador: Hélio Garcia Leite.

Dissertação (mestrado) - Universidade Federal de Viçosa,
Departamento de Engenharia Florestal, 2022.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2022.590>

Modo de acesso: World Wide Web.

1. Dendezeiro - Cultivo - Amazônia. 2. Inteligência artificial. 3. Dendezeiro - Rendimento - Métodos de simulação.
I. Leite, Hélio Garcia, 1960-. II. Universidade Federal de Viçosa.
Departamento de Engenharia Florestal. Programa de Pós-Graduação em Ciência Florestal. III. Título.

GFDC adapt. CDD 22. ed. 634.92831

Bibliotecário(a) responsável: Alice Regina Pinto Pires CRB-6/2523

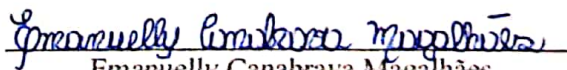
EMANUELLY CANABRAVA MAGALHÃES


MODELAGEM DA PRODUÇÃO DE *Elaeis guineensis* Jacq. UTILIZANDO REDES
NEURONAIS ARTIFICIAIS

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

APROVADA: 22 de agosto de 2022.

Assentimento:


Emanuella Canabrava Magalhães
Autora


Helio Garcia Leite
Orientador

À Mainha e Painho.

AGRADECIMENTOS

Agradeço a Deus e ao Universo por colocar na minha vida as melhores oportunidades e me guiar por caminhos de luz e amor.

À Mainha e Painho que lá do Norte de Minas mandam todo o amor do mundo, sonham junto comigo e torcem sempre pelo meu sucesso.

Aos meus irmãos Nanda e Gu que são meu refúgio pra rir das coisas mais fúteis da vida.

Aos meus gatinhos/filhos, Simba e Yá, que morro de saudades e sempre que precisava de um momento de fofura do dia tinha à disposição fotos, vídeos, além das ligações.

Aos meus melhores amigos Cris, Paloma e Saulo que em horas de ligação ou videochamadas transformam um dia ruim no melhor dia possível.

À Jianne e a Kemely, amigas que fiz na universidade e que foram e são pessoas essenciais pra eu continuar em Viçosa.

Aos meus orientadores Helio, Carlos Araújo e Carlos Pedro pela paciência, cuidado, disposição, ensinamentos, sabedoria, experiência, conselhos e por compartilharem sua inteligência e mente cheia de ideias.

Aos membros da banca, professores Júlio Neves e Haroldo Paiva pela disposição em participar dessa etapa e em compartilhar suas ideias e conhecimentos para agregar à pesquisa.

À Agropalma, especialmente ao Rodrigo Sales, Carlos Garcia e Juliana Machado pela disponibilidade em esclarecer dúvidas a todo momento, pelas reuniões mensais remotas que fizeram toda a diferença, pela viagem oferecida, na qual pude conhecer um pouco da cadeia produtiva da empresa e aprender com as pessoas ali presentes, por todo o cuidado e confiança depositada no nosso projeto.

À Viçosa por me acolher.

À Universidade Federal de Viçosa pela oportunidade de fazer parte de sua história.

Ao Departamento de Engenharia Florestal por ser uma segunda casa.

Ao Programa de Pós-Graduação em Ciência Florestal pela oportunidade de fazer pesquisa e contribuir pro avanço científico do Brasil.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudos.

À Vida.

Obrigada a todos que fizeram e fazem parte da minha história.

“Nem todo mundo vai compreender isso tudo que você é, o que não significa que você deva se esconder ou se calar. O mundo tem medo de mulheres extraordinárias”.

(Ryane Leão)

RESUMO

MAGALHÃES, Emanuely Canabrava, M.Sc., Universidade Federal de Viçosa, agosto de 2022. **Modelagem da produção de *Elaeis guineensis* Jacq. utilizando redes neurais artificiais.** Orientador: Helio Garcia Leite. Coorientadores: Carlos Alberto Araújo Júnior e Carlos Pedro Boechat Soares.

A crescente demanda pelo óleo de palma exige a produção em larga escala da cultura e uso otimizado dos recursos ambientais relacionados ao crescimento e desenvolvimento da espécie. O rendimento do óleo está diretamente associado à produção de cachos destinados à fábrica, por isso adotar ferramentas computacionais que possibilitem estimativas confiáveis da produção de cachos é tarefa primordial e importante para direcionar tomadas de decisão. Redes neurais artificiais (RNAs) correspondem a uma das técnicas de inteligência artificial de amplo uso e eficácia comprovada em estudos de modelagem da produção florestal, principalmente, produtos madeireiros, no entanto para o óleo de palma ainda não é uma realidade. Em virtude disso, o presente estudo teve como objetivo avaliar se redes neurais artificiais são capazes de estimar a produção mensal corrente e futura de cachos da palma-de-óleo utilizando variáveis do inventário florestal (inflorescências e cachos), elementos climáticos (temperatura, precipitação, umidade relativa do ar e insolação), déficit hídrico, solo, assim como variáveis do cadastro e gestão dos plantios (Fazenda, Ano de plantio, Material genético). Para a estimativa mensal corrente foram utilizadas as variáveis do inventário do mês atual e variáveis climáticas referentes ao mês anterior. A estimativa futura inclui a previsão da produção para os cinco meses futuros utilizando informações de produção e do inventário dos cinco meses anteriores ao mês a ser estimado. A configuração da RNA seguiu o mesmo padrão para ambas as abordagens, sendo uma parametrização simples com função de ativação sigmoide, algoritmo de treinamento resilient propagation, oito neurônios na camada oculta, e critério de parada de 3000 ciclos. A RNA estimou a produção corrente a nível de parcela com uma correlação acima de 0,6 e erro relativo percentual médio em torno de 13%. As variáveis de maior contribuição na modelagem incluíram àquelas relacionadas à gestão, solos, material genético e a contabilização de cachos maduros. As variáveis climáticas não exerceram tanta importância, porém devido a influência do elemento climático na produtividade do dendê é necessário mantê-las na modelagem. Na estimativa futura foi gerada correlação acima de 0,7 e o erro relativo percentual médio para a produção total estimada em cada mês futuro não ultrapassou 3%. Estes resultados confirmam o quanto a composição histórica do

comportamento produtivo é importante para a previsão de cachos. Em ambas as abordagens, a rede neuronal artificial demonstrou ser capaz de modelar a produção da palma-de-óleo, caracterizada por elevada variabilidade, abrindo oportunidades para realização de estudos futuros, combinando e utilizando novas variáveis para melhoria da acurácia das estimativas usando esta ferramenta.

Palavras-chave: Palma-de-óleo – Amazônia Oriental. Palma-de-óleo – inteligência artificial. Dendê. Palma-de-óleo – produtividade.

ABSTRACT

MAGALHÃES, Emanuely Canabrava, M.Sc., Universidade Federal de Viçosa, August, 2022. **Yield modeling of *Elaeis guineensis* Jacq. using artificial neural networks.** Adviser: Helio Garcia Leite. Co-Advisers: Carlos Alberto Araújo Júnior and Carlos Pedro Boechat Soares.

The growing demand for palm oil requires large-scale production and optimal use of environmental resources related to the growth and development of the species. Oil yield is directly associated with the bunches yield destined for the factory, so adopting computational tools that allow reliable estimates of bunch yield is a primordial and important task to guide decision making. Artificial Neural Networks (ANNs) correspond to one of the widely used and effective artificial intelligence techniques in studies of modeling forest production, mainly wood products, however for oil palm is not yet a reality. The present study aimed to evaluate whether artificial neural networks are able to estimate the current and future monthly yield of oil palm bunches, using forest inventory variables (inflorescences and bunches), weather elements (temperature, precipitation, humidity relative and insolation), water deficit, soil, as well as variables of the register and management of plantations (Farm, Year of planting, Genetic material). For the current monthly estimate, the inventory variables of the current month and climatic variables referring to the previous month were used. Future estimation includes forecasting yield for the future five months using yield and inventory information for the five months prior to the month to be estimated. The ANN estimated the current yield at the plot level with a correlation above 0.6 and average percentage relative error around 13%. The variables of greatest contribution in the modeling included those related to management, soils, genetic material and the accounting of mature bunches. The climatic variables did not exert much importance, however, due to the influence of the climatic element on oil palm productivity, it is necessary to keep them in the modeling. In the future estimate, a correlation above 0.7 was generated and the average percentage relative error for the estimated total yield in each future month did not exceed 3%. These results confirm how important the historical composition of the productive behavior is for the prediction of bunches. In both approaches, the artificial neural network proved to be able to model the oil palm yield, characterized by high variability, opening opportunities for future studies, combining and using new variables to improve the accuracy of estimates using this tool.

Keywords: Oil palm – Eastern Amazon. Oil palm – artificial intelligence. Oil palm. Oil palm – productivity.

SUMÁRIO

1	INTRODUÇÃO GERAL	13
2	CAPÍTULO 1 – FUNDAMENTAÇÃO TEÓRICA.....	15
	2.1 Caracterização da espécie	15
	2.2 Exigências ambientais	16
	2.3 Importância econômica	18
	2.4 Modelagem da produção	19
	2.5 Redes Neurais Artificiais.....	20
	2.6 Referências bibliográficas	22
3	CAPÍTULO 2 – PREDIÇÃO DA PRODUÇÃO DE CACHOS DA PALMA-DE-ÓLEO UTILIZANDO REDE NEURONAL ARTIFICIAL	29
	3.1 Introdução.....	29
	3.2 Material e métodos	30
	3.2.1 Área de estudo	30
	3.2.2 Caracterização dos plantios e organização da base de dados	31
	3.2.3 Predição da produção de cachos da palma-de-óleo	32
	3.2.4 Avaliação das redes neuronais artificiais.....	33
	3.2.5 Determinação das variáveis de maior contribuição na modelagem	34
	3.3 Resultados e discussão	35
	3.3.1 Predição da produção de cachos.....	36
	3.3.2 Variáveis de maior contribuição na modelagem	43
	3.4 Conclusão	48
	3.5 Referências bibliográficas	48
4	CAPÍTULO 3 – PROJEÇÃO DA PRODUÇÃO DE CACHOS DA PALMA-DE-ÓLEO UTILIZANDO REDE NEURONAL ARTIFICIAL	54

4.1	Introdução	54
4.2	Material e métodos	55
4.2.1	Área de estudo	55
4.2.2	Organização da base de dados	56
4.2.3	Projeção da produção de cachos da palma-de-óleo	57
4.2.4	Avaliação das redes neuronais artificiais.....	57
4.3	Resultados e discussão	58
4.3.1	Projeção da produção de cachos da palma-de-óleo	58
4.4	Conclusão	72
4.5	Referências bibliográficas	72
5	CONCLUSÃO GERAL	76

1 INTRODUÇÃO GERAL

A palma-de-óleo ou dendê, palmeira originária da Costa Ocidental Africana, tem ganhado destaque no cenário econômico mundial, em função da forte participação nas indústrias alimentícia, medicinal e cosmética, bem como da crescente necessidade de uso na área energética, devido ao aumento dos preços do petróleo bruto, recursos fósseis limitados e constantes preocupações com as mudanças climáticas, que exigem a substituição de combustíveis fósseis por biocombustíveis (BOARI, 2008; DUMRONGSIRI, 2012; JOHARI et al., 2015; DEY et al., 2021). Além disso, apresenta alto rendimento do óleo quando comparado a culturas como soja, mamona e amendoim, sendo o óleo vegetal mais produzido e comercializado no mundo, com 88% da produção concentrada no continente asiático (FAS, 2022).

No Brasil, a espécie teve seu cultivo impulsionado na Amazônia em função do material genético superior, disponibilidade de terras e condições climáticas favoráveis, sendo o estado do Pará o maior produtor nacional, concentrando 82% da produção (CÉSAR; BATALHA, 2013; VILLELA et al., 2014; NAHUM et al., 2020).

Atender a elevada demanda do óleo de palma não é tarefa simples, pois a produção em larga escala aumenta os custos, gera necessidade de grandes extensões de terra para cultivo, além da taxa de rendimento do produto ser dependente de variáveis climáticas, edáficas e operacionais, trazendo maior dificuldade ao processo de tomada de decisão (ZARRILLI, 2006; DUMRONGSIRI, 2012; SYAFIIE; VALIZADEH, 2016). Por isso, a construção de um plano de manejo é essencial, visto a necessidade de se obter os produtos certos, nas quantidades certas, no lugar e hora certos e com custo adequado, de forma a atender à solicitação do cliente (SEMBIRING et al., 2017).

O planejamento da produção do óleo de palma é dependente da modelagem da produção, etapa fundamental para a tomada de decisão em um empreendimento. A modelagem é feita usualmente empregando modelos de regressão. Nesse caso, é gerado um modelo estatístico que representa a relação entre a produtividade e as variáveis que a influenciam, para predição do crescimento e produção da espécie (BISWAS; NARAHARI, 2004; BRIASSOULIS, 2020).

Uma alternativa ao emprego de regressão é o uso de técnicas de inteligência artificial (IA), especificamente o aprendizado de máquina, como redes neurais artificiais, lógica fuzzy,

sistemas neuro-fuzzy, árvore de decisão, support vetor machine, dentre outras. Estas, têm sido aplicadas em estudos de modelagem da produção, sendo capazes de gerar respostas muito próximas da realidade (KARTIKA et al., 2016; CHAPMAN et al., 2018).

O emprego de métodos de aprendizado de máquina tem sido cada vez mais frequente no setor florestal para diversas aplicações (GÖRGENS et al., 2015; PÉRIÉ; BLOIS, DE, 2016; LIU et al., 2018; DIAMANTOPOULOU et al., 2018; DINIZ et al., 2021), destacando a modelagem do crescimento e da produção de madeira (ZHANG et al., 2000; BINOTI, 2010; CASTRO et al., 2013; BINOTI et al., 2015; DIAMANTOPOULOU et al., 2015; VIEIRA et al., 2018; CHIARELLO et al., 2019).

Considerando a eficiência já comprovada do uso de técnicas de inteligência artificial para modelagem de crescimento e produção de madeira e a importância da prognose de palma-de-óleo, o presente estudo objetiva avaliar a eficiência do emprego de redes neurais artificiais para predição e projeção da produção de cachos da palma-de-óleo no estado do Pará.

2 CAPÍTULO 1 – FUNDAMENTAÇÃO TEÓRICA

2.1 Caracterização da espécie

A palma-de-óleo ou dendê é uma palmeira originária da Costa Ocidental Africana, pertencente ao grupo das monocotiledôneas, à família botânica Arecaceae, subfamília Areicoideae, e gênero *Elaeis*, que compreende duas principais espécies: *Elaeis guineensis* Jacq. e *Elaeis oleifera* nativas dos continentes africano e americano, respectivamente (TELES, 2014; BARCELOS et al., 2015; CORLEY; TINKER, 2016).

Possui entre 15 e 18 m de altura, podendo atingir 30 m em plantios densos. As folhas são do tipo pinada e o caule é colunar solitário, grosso e ereto na espécie africana e pouco inclinado na americana. O cacho apresenta formato ovoide, podendo alcançar 50 cm de comprimento e 35 cm de largura. Aos 3 anos de idade, as palmeiras produzem cachos pesando até 5 kg, enquanto aos 15 anos de idade o peso ultrapassa os 25 kg, dependendo das condições ambientais e do material genético. Cada cacho produz em média 1500 frutos (BORGES et al., 2016; CORLEY; TINKER, 2016).

O fruto é do tipo drupa séssil, mais ou menos esférica e alongada com o topo protuberante (Figura 1). Quanto ao comprimento pode variar de 2 a 5 centímetros e pesar entre 3 a 30 gramas. Ao cortar o fruto podem ser visualizados 4 componentes: Exocarpo, dito como a casca do fruto; Mesocarpo, de coloração amarela ou alaranjada e de onde o óleo é extraído; Endocarpo, de cor negra, muito duro e envolvido por fibras aderentes e o Endosperma de formato ovoide, ocupando toda a cavidade do endocarpo (SILVA, 2006).



Figura 1 – Fruto da Palma-de-óleo

Fonte: O autor (2022)

São definidas três variedades da planta conforme espessura da polpa: Dura, quando o fruto apresenta endocarpo espesso com pouca polpa; Psífera, onde o fruto não apresenta endocarpo e geralmente é abortivo e Tenera, com frutos de endocarpo fino, maior proporção de polpa e alta produção de óleo (FERREIRA et al., 2012).

É uma espécie monoica, na qual inflorescências femininas e masculinas estão distribuídas no mesmo indivíduo e são produzidas em ciclos alternados de duração variável dependentes de fatores genéticos, idade e condições ambientais. A maturidade das inflorescências é atingida 32 a 36 meses após a germinação das sementes (ADAM et al., 2005; CORLEY; TINKER, 2016).

A cultura é perene com ciclo produtivo de 25 a 30 anos, sendo a produção do fruto iniciada no terceiro ano, podendo atingir 25 toneladas de cacho por hectare (SIMAS; CAMARGO PENTEADO, 2019).

O principal produto da espécie é o óleo, sendo sua produção dependente da composição física do cacho, como o percentual de frutos no cacho, percentual de mesocarpo e de óleo no fruto, que afetam diretamente a taxa de extração industrial (TEI) (SILVA, 2006, TELES, 2014). São extraídos o óleo de palma do mesocarpo (polpa fibrosa) e o óleo de palmiste da semente, havendo um rendimento de 4 a 5 toneladas de óleo de palma, o que é bastante elevado quando comparados com culturas como soja, mamona e amendoim (MURPHY, 2014; SIMAS; CAMARGO PENTEADO, 2019).

2.2 Exigências ambientais

A produtividade da espécie é dependente de condições climáticas específicas, por isso se desenvolve melhor em regiões tropicais. Os elementos climáticos de maior influência são temperatura, precipitação pluviométrica, umidade relativa do ar e radiação solar, além destes a deficiência hídrica anual é outro fator limitante da produção (BOARI et al., 2010; BARCELOS et al., 2015; CORLEY; TINKER, 2016; RHEBERGEN et al., 2016).

A temperatura média anual ideal deve variar entre 24 e 28°C, precipitações pluviométricas anuais atingindo uma média superior a 2000 mm, com valores mensais mínimos

de 100 mm, chuvas bem distribuídas ao longo do ano, sem ocorrência de estações secas definidas. As variações pluviométricas anuais interferem diretamente na diferenciação sexual das inflorescências e na produção dos cachos em um intervalo de até 28 meses (CARR, 2011; FARIAS, 2014; CORLEY; TINKER, 2016).

A espécie é sensível ao estresse hídrico, visto que a disponibilidade constante de água no solo é importante para o desenvolvimento da palmeira. O déficit hídrico médio anual não deve ultrapassar 200 mm, além disso a ocorrência de mais de 3 meses secos consecutivos afeta a produtividade, isto porque contribui para redução do aparecimento de folhas novas, maior quantidade de inflorescência masculina e diminuição do peso dos cachos (BOARI et al., 2010; BARCELOS et al., 2015; CORLEY; TINKER, 2016; RHEBERGEN et al., 2016).

A umidade relativa do ar influencia a capacidade fotossintética da espécie, sendo ideal uma variação mensal entre 75 e 90%. O cultivo em condições de baixa umidade relativa do ar restringe a abertura estomática e absorção de CO₂ pela planta (BOARI et al., 2010; BARCELOS et al., 2015).

Elevados níveis de radiação solar é crucial para a produção de cachos, em condições de sombreamento a taxa de crescimento é reduzida em qualquer idade do plantio, já que provoca a redução da produção de inflorescências femininas. A taxa de CO₂ absorvido pela folha verde durante a fotossíntese é favorecida com a radiação incidente. Portanto é necessária uma disponibilidade de luz solar de pelo menos 5 horas diárias ou 1600 horas no ano (BOARI et al., 2010; BARCELOS et al., 2015; CORLEY; TINKER, 2016).

Adapta-se a diferentes tipos de solos, apresentando elevada produtividade em solos com alta fertilidade natural, desde que as condições climáticas sejam atendidas. Pode ser cultivada em solos pobres e ácidos sob nutrição adequada (BOARI et al., 2010; CORLEY; TINKER, 2016; RHEBERGEN et al., 2016).

Locais com relevo plano a suave ondulado são os mais recomendados. Enquanto topografias acidentadas acima de 5% exigem a construção de terraços para evitar erosão. Em solos com declividade entre 2 e 5%, plantios em curvas de nível são mais indicados, inclinações do terreno superiores a 10% devem ser evitadas e em inclinações abaixo de 5% são indicadas plantas de cobertura. Quanto à estrutura do solo, observa-se que solos profundos, sem camadas de impedimento e boa permeabilidade possibilitam bom desenvolvimento da palma-de-óleo (BOARI et al., 2010).

2.3 Importância econômica

Dentre os óleos de origem vegetal, o óleo de palma é o mais produzido e comercializado no mundo, destacando-se nas indústrias alimentícia, cosmética, farmacêutica e energética, em função do alto desempenho agrícola e industrial (MBA et al., 2015; DEY et al., 2021).

A maior aplicação do óleo encontra-se na indústria alimentícia, utilizado como matéria-prima na fabricação de produtos como margarina, maionese, biscoitos, sorvetes, dentre outros (MBA et al., 2015; TEH et al., 2017). O fruto da palma também é processado e dele são gerados subprodutos como a torta de palmiste, engaços (cachos vazios), fibras, cascas e efluentes líquidos, usados como fonte de nutrientes e energia (MOHAMMAD et al., 2012; BARTHEL et al., 2018; GONZALEZ-DIAZ et al., 2021; SUBRAMANIAM et al., 2021).

Nos últimos anos, devido às constantes preocupações com as mudanças climáticas, aumento dos preços do petróleo bruto e recursos fósseis limitados, nota-se uma tendência de substituição dos combustíveis fósseis por fontes de energia renováveis, os biocombustíveis, de forma a reduzir a emissão de gases de efeito estufa na atmosfera (ZARRILLI, 2006; BOARI, 2008; DUMRONGSIRI, 2012; JOHARI et al., 2015; SYAFIIE; VALIZADEH, 2016; DEY et al., 2021). Tal movimento, possibilitou o uso do óleo de palma na indústria energética para produção do biodiesel, que apresenta como vantagens: alta produtividade por unidade de área, elevado balanço energético e semelhança com o diesel convencional. No entanto, o custo de produção, especialmente da mão-de-obra é elevado (MINISTÉRIO DA AGRICULTURA PECUÁRIA E ABASTECIMENTO, 2018).

Os principais produtores do óleo de palma estão localizados no continente asiático: Indonésia (58%), Malásia (26%), e Tailândia (4%), somando 88% da produção mundial. No continente americano, destacam-se a Colômbia, Equador, Brasil, Honduras, Costa Rica e Guatemala e; no continente africano, Nova Guiné e Costa do Marfim (FAS, 2022).

No Brasil, a espécie foi introduzida no século 16, com cultivo impulsionado na Amazônia em função do material genético superior, disponibilidade de terras e condições climáticas favoráveis. O país produz anualmente cerca de 550 mil toneladas de óleo de palma, sendo 82% da produção concentrada no estado do Pará, maior produtor nacional (FAS, 2022). Na região, o cultivo do dendezeiro foi incentivado pela ação estatal, para desenvolvimento econômico do espaço rural, o que promoveu a integração entre o agricultor familiar e a agroindústria, gerando emprego, renda e inclusão social. A chegada de empreendimentos no

ambiente rural proporcionou melhoria na estrutura local, referente à qualidade das estradas, transporte, telefonia e eletricidade (CÉSAR; BATALHA, 2013; VILLELA et al., 2014; BARTHEL et al., 2018; NAHUM et al., 2020).

Apesar das condições edafoclimáticas favoráveis e disponibilidade de terras para cultivo, a produção do óleo de palma no Brasil, corresponde a apenas 1% da produção mundial, além de não atender a demanda interna, o que implica na importação do produto (MONTEIRO et al., 2014; MINISTÉRIO DA AGRICULTURA PECUÁRIA E ABASTECIMENTO, 2018), tornando-se necessário manejar os plantios de forma a otimizar a produção e, conseqüentemente, atender a demanda.

A expansão do cultivo de dendezeiros deve ser feita de forma responsável e sustentável, em função dos possíveis impactos ambientais. Barthel et al. (2018) citam como principal impacto, o desmatamento de florestas naturais, e, conseqüente, perda da biodiversidade e emissão de gases de efeito estufa, já que sua capacidade de armazenamento de carbono por hectare é muito inferior a armazenada por florestas tropicais.

2.4 Modelagem da produção

Com o crescimento da indústria do óleo de palma e sua maior relevância no cenário econômico mundial, torna-se importante entender a relação entre a produtividade e as variáveis que exercem influência sobre a mesma (KARTIKA et al., 2016). Para tanto, a modelagem é uma técnica que vem sendo utilizada para melhor compreensão dos processos florestais, de forma a auxiliar a tomada de decisão. Modelar um fenômeno, significa simplificar o ambiente real em um ambiente computacional, a partir de formulações matemáticas que o represente (BISWAS; NARAHARI, 2004; BRIASSOULIS, 2020).

Modelos de regressão linear e não linear vêm sendo amplamente utilizados em estudos de modelagem (MOREIRA et al., 2015; SENA et al., 2015; TÉO et al., 2017; BAPTISTA et al., 2019). No entanto, à medida que a complexidade das relações entre as variáveis aumenta, a eficiência dos modelos diminui e surge a necessidade de uso de outros métodos de predição (BINOTI et al., 2013).

Técnicas de inteligência artificial (IA) têm ganhado destaque e sendo aplicadas como alternativa aos métodos tradicionais, devido à maior capacidade de generalização e menor sensibilidade a ruídos e outliers (BAYAT et al., 2020). Redes neuronais artificiais (RNA), lógica fuzzy, sistema neuro-fuzzy e support vector machine (SVM) são algumas técnicas de IA encontradas em estudos da área florestal (ZHANG et al., 2000; BINOTI, 2010; BINOTI et al., 2013; CASTRO et al., 2013; ÖZÇELIK et al., 2013; GÖRGENS et al., 2015; DIAMANTOPOULOU et al., 2015, 2018; PÉRIÉ; BLOIS, 2016; LIU et al., 2018; VIEIRA et al., 2018; CHIARELLO et al., 2019; BAYAT et al., 2020; DINIZ et al., 2021).

Na literatura são encontrados alguns estudos de modelagem para compreensão de fenômenos envolvendo a palma-de-óleo, como em Kartika et al. (2016), que utilizaram redes neuronais artificiais para previsão do rendimento do óleo de palma a partir de variáveis climáticas; Chapman et al. (2018) predisseram a produção de frutos do dendê, do peso médio dos cachos e do número médio de cachos por hectare a partir de variáveis como precipitação, fertilização e idade da árvore, adotando como ferramentas de processamento as Redes bayesianas e Redes neuronais artificiais; Iguodala et al. (2016), aplicaram um modelo misto para avaliar a influência de potássio e magnésio na produção de cachos de frutos da palma-de-óleo; Hoffmann et al. (2014), desenvolveram um modelo (PALMSIM) que simula mensalmente o crescimento e rendimento potencial do dendê conforme radiação solar; e Plazas-González et al. (2018), modelaram o mecanismo de hidrotreamento do óleo de palma para determinar as condições ótimas de produção do diesel verde. Alguns outros estudos vêm utilizando RNAs para classificar a maturação do cacho da palma-de-óleo a partir de imagens (ALFATNI et al., 2018, 2020), bem como para previsão do tempo ótimo de processamento do fruto na fábrica (ADIZUE et al., 2020).

2.5 Redes Neuronais Artificiais

Inspirados no funcionamento do cérebro humano, McCulloch e Pitts (1943) desenvolveram o primeiro modelo matemático de uma rede neuronal, ao simplificar um neurônio biológico em um neurônio artificial (Figura 2). O neurônio artificial é constituído por n entradas e uma saída, as entradas são associadas a pesos positivos ou negativos. Calcula-se a soma ponderada dos valores recebidos e respectivos pesos, caso o valor seja superior ao limiar

do neurônio, é ativada a função de ativação que gera a saída (BRAGA et al., 2007; BINOTI, 2010; AN et al., 2022).

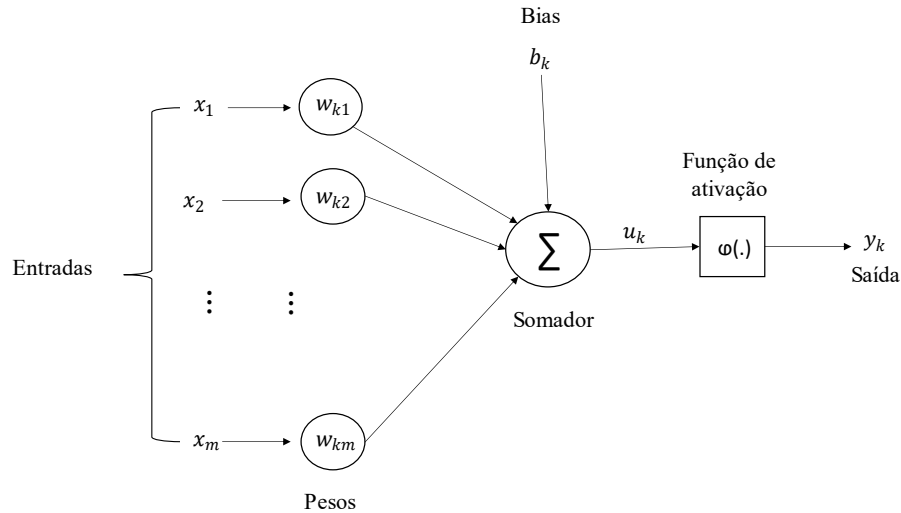


Figura 2 - Representação do neurônio artificial

Fonte: Adaptado de Haykin (2009)

As redes neuronais artificiais (RNAs) são sistemas computacionais constituídos por várias unidades de processamento denominadas neurônios artificiais, os quais estão conectados entre si para executar determinada tarefa (LIAO et al., 2012; MUMALI, 2022). Os neurônios artificiais se assemelham ao neurônio biológico em dois aspectos: a RNA adquire conhecimento a partir de um processo de aprendizagem e a conexão entre os neurônios, denominados pesos sinápticos, são responsáveis por armazenar o conhecimento adquirido (HAYKIN, 2009; BINOTI et al., 2013).

A RNA é constituída por três camadas: a camada de entrada, onde são recebidas as informações, camada de saída, responsável por gerar as saídas/respostas e a camada oculta que conecta ambas as camadas anteriores (HAYKIN, 2009; LIAO et al., 2012) (Figura 3). O treinamento consiste em apresentar um conjunto específico de exemplos referentes ao problema de estudo, chamados de dados de entrada, bem como o padrão desejado de respostas, ditos como dados de saída. Os dados de entrada são repetidamente apresentados à RNA e os pesos atualizados conforme novos dados são inseridos, o que constitui o aprendizado da RNA (LIAO et al., 2012).

Trabalhar com uma rede neuronal exige uma configuração base para início do processo de aprendizagem que inclui: 1. Número de neurônios na camada oculta, 2. Algoritmo de aprendizado, cuja função é modificar os pesos sinápticos de forma a atingir determinado objetivo, 3. Função de ativação, responsável por decidir se um neurônio deve ser ativado ou não, para que pequenas alterações nos pesos e bias resultem em pequenas alterações na saída, 4. Critério de parada do treinamento, definido como número de ciclos ou erro médio. (HAYKIN, 2009).

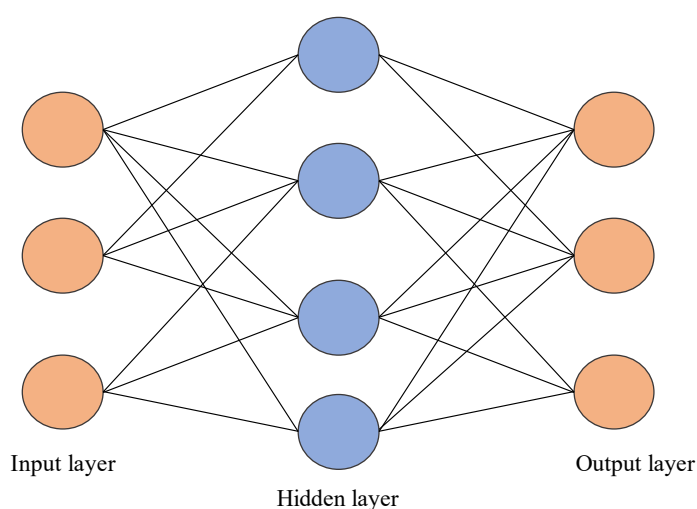


Figura 3 – Estrutura de uma rede neuronal artificial

2.6 Referências bibliográficas

ADAM, H.; JOUANNIC, S.; ESCOUTE, J.; et al. Reproductive developmental complexity in the African oil palm (*Elaeis guineensis*, Arecaceae). **American Journal of Botany**, v. 92, n. 11, p. 1836–1852, 2005.

ADIZUE, U. L.; NWANYA, S. C.; OZOR, P. A. Artificial neural network application to a process time planning problem for palm oil production. **Engineering and Applied Science Research**, v. 47, n. 2, p. 161–169, 2020.

ALFATNI, M. S. M.; MOHAMED SHARIFF, A. R.; BEJO, S. K.; SAAED, O. M. BEN; MUSTAPHA, A. Real-time oil palm FFB ripeness grading system based on ANN, KNN and

SVM classifiers. **IOP Conference Series: Earth and Environmental Science**, v. 169, n. 1, 2018.

ALFATNI, M. S. M.; MOHAMED SHARIFF, A. R.; SAAED, O. M. BEN; ALBHBAH, A. M.; MUSTAPHA, A. Colour Feature Extraction Techniques for Real Time System of Oil Palm Fresh Fruit Bunch Maturity Grading. **IOP Conference Series: Earth and Environmental Science**, v. 540, n. 1, 2020.

AN, J.; LIU, F.; SHEN, F.; et al. IC neuron: An efficient unit to construct neural networks. **Neural Networks**, v. 145, p. 177–188, 2022. Elsevier Ltd. Disponível em: <<https://doi.org/10.1016/j.neunet.2021.10.005>>.

BAPTISTA, F. R. C. B.; ATAÍDE, D. H. S.; ROCHA, P. V.; et al. Stratification in the adjustment of hypsometric relationship models in forest fragment. **Scientific Electronic Archives**, v. 12, n. 6, p. 52, 2019.

BARCELOS, E.; ALMEIDA RIOS, S. DE; CUNHA, R. N. V.; et al. Oil palm natural diversity and the potential for yield improvement. **Frontiers in Plant Science**, v. 6, n. MAR, p. 1–16, 2015.

BARTHEL, M.; JENNINGS, S.; SCHREIBER, W.; SHEANE, R.; ROYSTON, S. **Study on the environmental impact of palm oil consumption and on existing sustainability standards**. 2018.

BAYAT, M.; BETTINGER, P.; HEIDARI, S.; et al. Estimation of tree heights in an uneven-aged, mixed forest in northern Iran using artificial intelligence and empirical models. **Forests**, v. 11, n. 3, 2020.

BINOTI, M. L. M. DA S. **Redes neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto**, 2010. Dissertação (Mestrado), Viçosa: Universidade Federal de Viçosa.

BINOTI, M. L. M. DA 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, 2013.

BINOTI, M. L. M. DA S.; LEITE, H. G.; BINOTI, D. H. B.; GLERIANI, J. M. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **Cerne**, v. 21, n. 1, p. 97–105, 2015. Federal University of Lavras.

BISWAS, S.; NARAHARI, Y. Object oriented modeling and decision support for supply chains. **European Journal of Forest Research**, v. 153, p. 704–726, 2004.

- BOARI, A. J. **Estudos realizados sobre o amarelecimento fatal do Dendezeiro (*Elaeis guineensis* Jacq.) no Brasil**. Belém: Embrapa Amazônia Oriental, 2008.
- BOARI, A. J.; BARROS, A. H.; GONÇALVES, A. O.; et al. **Zoneamento agroecológico, produção e manejo para a cultura da Palma de Óleo na Amazônia**. 1st ed. Rio de Janeiro: Agência 2A comunicação, 2010.
- BORGES, A. DE J.; COLLICCHIO, E.; CAMPOS, G. A. A cultura da palma de óleo (*Elaeis guineensis* Jacq.) no Brasil e no mundo: aspectos agronômicos e tecnológicos - uma revisão. **Revista Liberato**, v. 17, n. 27, p. 65–78, 2016.
- BRAGA, A. DE P.; CARVALHO, A. C. P. DE L. F. DE; LUDERMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações**. 2nd ed. Rio de Janeiro: LTC Editora, 2007.
- BRIASSOULIS, H. **The Web Book of Regional Science Sponsored by Analysis of Land Use Change: Theoretical and modeling approaches by Helen Briassoulis**. 2nd ed. WVU Research Repository, 2020.
- CARR, M. K. V. THE water relations and irrigation requirements of oil palm (*Elaeis guineensis*): A review. **Experimental Agriculture**, v. 47, n. 4, p. 629–652, 2011.
- CASTRO, R. V. O.; SOARES, C. P. B.; LEITE, H. G.; et al. Individual Growth Model for Eucalyptus Stands in Brazil Using Artificial Neural Network. **ISRN Forestry**, v. 2013, p. 1–12, 2013.
- CÉSAR, A. DA S.; BATALHA, M. O. Brazilian biodiesel: The case of the palm's social projects. **Energy Policy**, v. 56, p. 165–174, 2013.
- CHAPMAN, R.; COOK, S.; DONOUGH, C.; et al. Using Bayesian networks to predict future yield functions with data from commercial oil palm plantations: A proof of concept analysis. **Computers and Electronics in Agriculture**, v. 151, n. October 2017, p. 338–348, 2018.
- CHIARELLO, F.; STEINER, M. T. A.; OLIVEIRA, E. B. DE; ARCE, J. E.; FERREIRA, J. C. Artificial neural networks applied in forest biometrics and modeling: State of the art (january/2007 to july/2018). **Cerne**, v. 25, n. 2, p. 140–155, 2019.
- CORLEY, R. H. V; TINKER, P. B. **The Oil Palm**. BlackWell Science Ltd, 2016.
- DEY, S.; REANG, N. M.; DAS, P. K.; DEB, M. A comprehensive study on prospects of economy, environment, and efficiency of palm oil biodiesel as a renewable fuel. **Journal of Cleaner Production**, v. 286, p. 124981, 2021. Elsevier Ltd. Disponível em: <<https://doi.org/10.1016/j.jclepro.2020.124981>>.

DIAMANTOPOULOU, M. J.; ÖZÇELİK, R.; CRECENTE-CAMPO, F.; ELER, Ü. Estimation of Weibull function parameters for modelling tree diameter distribution using least squares and artificial neural networks methods. **Biosystems Engineering**, v. 133, p. 33–45, 2015.

DIAMANTOPOULOU, M. J.; ÖZÇELİK, R.; YAVUZ, H. Tree-bark volume prediction via machine learning: A case study based on black alder's tree-bark production. **Computers and Electronics in Agriculture**, v. 151, p. 431–440, 2018. Elsevier B.V.

DINIZ, É. S.; LORENZON, A. S.; CASTRO, N. L. M. DE; et al. Forecasting frost risk in forest plantations by the combination of spatial data and machine learning algorithms. **Agricultural and Forest Meteorology**, v. 306, n. April, 2021.

DUMRONGSIRI, A. Mathematical model for production, logistics and plant capacity planning of oil palm bunches. **Mae Fah Luang University International Conference 2012**. Anais..., v. 1, p.37–39, 2012. Pathum Thani.

FARIAS, M. P. **Caracterização do desenvolvimento floral de *Elaeis oleifera* (H. B. K.) Cortés**, 2014. Dissertação (Mestrado), Lavras: Universidade Federal de Lavras – UFLA.

FAS - FOREIGN AGRICULTURAL SERVICES. Oil, Palm explorer. Disponível em: <<https://ipad.fas.usda.gov/cropexplorer/cropview/commodityView.aspx?cropid=4243000>>. Acesso em: 9/3/2022.

FERREIRA, C. B. B.; LOPES, M. T. G.; LOPES, R.; et al. Diversidade genética molecular de progênies de dendezeiro. **Pesquisa Agropecuária Brasileira**, v. 47, n. 3, p. 378–384, 2012.

GONZALEZ-DIAZ, A.; PATAQUIVA-MATEUS, A.; GARCÍA-NÚÑEZ, J. A. Recovery of palm phytonutrients as a potential market for the by-products generated by palm oil mills and refineries—A review. **Food Bioscience**, v. 41, n. February, 2021.

GÖRGENS, 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.

HAYKIN, S. **Neural Networks and Learning Machines**. 3rd ed. Ontario: Pearson, 2009.

HOFFMANN, M. P.; VERA, A. C.; WIJK, M. T. VAN; et al. Simulating potential growth and yield of oil palm (*Elaeis guineensis*) with PALMSIM: Model description, evaluation and application. **Agricultural Systems**, v. 131, p. 1–10, 2014. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.agsy.2014.07.006>>.

IGUODALA, E. A.; THOMAS, E. J.; EDOKPAYI, A; TITILOLA, O. A Mixed model Analysis of a fertilizer experiment on oil palm in Nigeria. *Agricultural Sciences*, v. 7, n. 8, p. 521 – 530, 2016.

JOHARI, A.; NYAKUMA, B. B.; MOHD NOR, S. H.; et al. The challenges and prospects of palm oil based biodiesel in Malaysia. **Energy**, v. 81, p. 255–261, 2015. Elsevier. Disponível em: <<http://dx.doi.org/10.1016/j.energy.2014.12.037>>.

KARTIKA, N. D.; ASTIKA, I. W.; SANTOSA, E. Oil palm yield forecasting based on weather variables using artificial neural network. **Indonesian Journal of Electrical Engineering and Computer Science**, v. 3, n. 3, p. 626–633, 2016.

LIAO, Z.; WANG, B.; XIA, X.; HANNAM, P. M. Environmental emergency decision support system based on Artificial Neural Network. **Safety Science**, v. 50, n. 1, p. 150–163, 2012. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.ssci.2011.07.014>>.

LIU, Z.; PENG, C.; WORK, T.; et al. Application of machine-learning methods in forest ecology: Recent progress and future challenges. **Environmental Reviews**, v. 26, n. 4, p. 339–350, 2018.

MBA, O. I.; DUMONT, M. J.; NGADI, M. Palm oil: Processing, characterization and utilization in the food industry - A review. **Food Bioscience**, v. 10, p. 26–41, 2015.

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

MINISTÉRIO DA AGRICULTURA PECUÁRIA E ABASTECIMENTO. **Diagnóstico da Produção Sustentável da Palma de Óleo**. 1st ed. Brasília: Assessoria de Comunicação e Eventos, 2018.

MOHAMMAD, N.; ALAM, M. Z.; KABBASHI, N. A.; AHSAN, A. Effective composting of oil palm industrial waste by filamentous fungi: A review. **Resources, Conservation and Recycling**, v. 58, p. 69–78, 2012. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.resconrec.2011.10.009>>.

MONTEIRO, K. F. G.; HOMMA, A. KINGO O. Diferentes sistemas de produção com palma de óleo (*Elaeis guineensis* jaq.) e a participação do brasil no cenário internacional. **Embrapa Amazônia Oriental**, 2014.

MOREIRA, M. F. B.; THIERSCH, C. R.; ANDRADE, M. G. DE; SCOLFORO, J. R. S. Estimativa da relação hipsométrica com modelos não lineares ajustados por métodos bayesianos empíricos. **Cerne**, v. 21, n. 3, p. 405–441, 2015.

MUMALI, F. Artificial neural network-based decision support systems in manufacturing processes: A systematic literature review. **Computers and Industrial Engineering**, v. 165, n. January, p. 107964, 2022. Elsevier Ltd. Disponível em: <<https://doi.org/10.1016/j.cie.2022.107964>>.

MURPHY, D. J. The future of oil palm as a major global crop: Opportunities and challenges. **Journal of Oil Palm Research**, v. 26, n. MAR, p. 1–24, 2014.

NAHUM, J. S.; SANTOS, L. S. DOS; SANTOS, C. B. DOS. **Formação da dendeicultura na amazônia paraense**. v. 19, p. 1–14, 2020.

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.

PÉRIÉ, C.; BLOIS, S. DE. Dominant forest tree species are potentially vulnerable to climate change over large portions of their range even at high latitudes. **PeerJ**, v. 2016, n. 7, 2016.

PLAZAS-GONZÁLEZ, M.; GUERRERO-FAJARDO, C. A.; SODRÉ, J. R. Modelling and simulation of hydrotreating of palm oil components to obtain green diesel. **Journal of Cleaner Production**, v. 184, p. 301–308, 2018.

RHEBERGEN, T.; FAIRHURST, T.; ZINGORE, S.; et al. Climate, soil and land-use based land suitability evaluation for oil palm production in Ghana. **European Journal of Agronomy**, v. 81, p. 1–14, 2016. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.eja.2016.08.004>>.

SEMBIRING, P.; MAWENKANG, H.; SADYADHARMA, H. Mixed Integer Linear Programming model for Crude Palm Oil Supply Chain Planning. **IOP Conference Series: Materials Science and Engineering**. Anais..., v. 300, p.2–7, 2017. Medan.

SENA, A. L. M.; NETO, A. J. DA S.; OLIVEIRA, G. M. V.; CALEGARIO, N. Modelos lineares e não lineares com uso de covariantes para relação hipsométrica de duas espécies de pinus tropicais. **Ciência Florestal**, v. 25, n. 4, p. 960–980, 2015.

SILVA, J. S. O. **Produtividade de óleo de palma na cultura do dendê na Amazônia Oriental: influência do clima e do material genético**, 2006. Dissertação (Mestrado), Viçosa: Universidade Federal de Viçosa - UFV.

SIMAS, J. P.; CAMARGO PENTEADO, C. L. Energia e sustentabilidade: Análise da viabilidade do cultivo de Dendê para o desenvolvimento regional endógeno da Amazônia. **Boletim de Geografia**, v. 37, n. 1, p. 184, 2019.

SUBRAMANIAM, V.; LOH, S. K.; AZIZ, A. A. GHG analysis of the production of crude palm oil considering the conversion of agricultural wastes to by-products. **Sustainable Production and Consumption**, v. 28, p. 1552–1564, 2021. Elsevier B.V. Disponível em: <<https://doi.org/10.1016/j.spc.2021.09.004>>.

SYAFIIE, S.; VALIZADEH, M. Optimal planning of a biofuel supply chain using a multi-criteria optimisation model Maryam Valizadeh. **International Journal Process Systems Engineering**, v. 4, n. 1, p. 1–18, 2016.

TEH, S. S.; ONG, A. S. H.; MAH, S. H. Recovery and Utilization of Palm Oil Mill Effluent Source as Value-Added Food Products. **Journal of Oleo Science**, v. 66, n. 11, p. 1183–1191, 2017.

TELES, D. A. A. **Características físicas e rendimento mensal em Óleo de cachos de duas cultivares de dendezeiro cultivadas, sob irrigação, no Cerrado do Distrito Federal**, 2014. Dissertação (Mestrado), Brasília: Universidade de Brasília - UNB.

TÉO, S. J.; MACHADO, S. DO A.; FILHO, A. F.; TOMÉ, M. Relação hipsométrica geral com atributos biológicos para povoamentos de *Pinus taeda* L. **Cerne**, v. 23, n. 4, p. 403–411, 2017.

VIEIRA, G. C.; MENDONÇA, A. R. DE; SILVA, G. F. DA; et al. Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. **Science of the Total Environment**, v. 619–620, p. 1473–1481, 2018. Elsevier B.V. Disponível em: <<https://doi.org/10.1016/j.scitotenv.2017.11.138>>.

VILLELA, A. A.; JACCOUD, D. B.; ROSA, L. P.; FREITAS, M. V. Status and prospects of oil palm in the Brazilian Amazon. **Biomass and Bioenergy**, v. 67, p. 270–278, 2014. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.biombioe.2014.05.005>>.

ZARRILLI, S. The emerging biofuels market: Regulatory, Trade and Development implications. **United Nations Conference on Trade and Development (UNCTAD)**. Anais..., p.1–48, 2006. New York.

ZHANG, Q. BIN; HEBDA, R. J.; ZHANG, Q. J.; ALFARO, R. I. Modeling tree-ring growth responses to climatic variables using Artificial Neural Networks. **Forest Science**, v. 46, n. 2, p. 229–239, 2000.

3 CAPÍTULO 2 – PREDIÇÃO DA PRODUÇÃO DE CACHOS DA PALMA-DE-ÓLEO UTILIZANDO REDE NEURONAL ARTIFICIAL

3.1 Introdução

O Brasil apresenta elevado potencial para se tornar um dos maiores produtores do óleo de palma, em função da disponibilidade de terras e condições climáticas favoráveis. No entanto, a realidade é outra, e a produção atual ainda é bastante inferior quando comparada aos maiores produtores mundiais e não consegue atender à crescente demanda interna (OIT, 2020; DROST et al., 2021; FAS, 2022).

Para atendimento da demanda é necessário otimizar a produção, de modo a obter aproveitamento máximo dos recursos disponíveis, o que exige conhecimento aprimorado dos fatores envolvidos no crescimento e desenvolvimento da espécie para que a tomada de decisão seja assertiva. Assim, é importante que o gestor florestal utilize ferramentas computacionais capazes de relacionar/modelar a produção conforme variáveis envolvidas no processo produtivo e que o auxilie na resolução de problemas (BRIASSOULIS, 2020).

Ferramentas de inteligência artificial, como redes neurais artificiais, capazes de modelar problemas de maior complexidade, têm ganhado destaque na modelagem de produtos florestais, havendo uma série de estudos para espécies como eucalipto, pinus e teca, com o intuito de estimar a altura e volume de madeira, principalmente (GORGENS et al., 2009; BINOTI, 2010; BINOTI et al., 2015; LOUREIRO, 2016). Para a palma-de-óleo, que gera um produto não-madeireiro, estudos nesse sentido ainda são escassos, sendo a maioria concentrados nos países de maior produção da espécie (HOFFMANN et al., 2014; IGUODALA et al., 2016; KARTIKA et al., 2016; CHAPMAN et al., 2018; ADIZUE et al., 2020; YOUSEFI D.B. et al., 2021; WATSON-HERNÁNDEZ et al., 2022).

Logo, são importantes pesquisas de âmbito nacional que impulsionem a modelagem do crescimento e desenvolvimento da espécie aplicando tais ferramentas. Pensando nisso, o presente estudo visa avaliar se redes neurais artificiais são capazes de estimar a produção corrente de cachos de palma-de-óleo utilizando variáveis do inventário e climáticas.

3.2 Material e métodos

3.2.1 Área de estudo

A área de estudo pertence à empresa Agropalma, localizada no estado do Pará, região norte do Brasil (Figura 1). Apresenta 107 mil hectares de terras, dos quais aproximadamente 39 mil hectares são destinados a plantios seminais comerciais de *Elaeis guineensis* (Jacq.) da variedade Tenera, distribuídos pelos municípios de Tailândia (sede), Tomé-Açu, Moju e Acará, 64 mil hectares ocupados por reservas florestais e demais áreas de infraestrutura (AGROPALMA, 2019).

O tipo de solo predominante na região é o Argissolo Amarelo distrófico, de profundidade variável, ácido, bem drenado e textura média a argilosa. Em algumas áreas, ocorre o Gleissolo háplico distrófico, pouco profundo, ácido, hidromórfico, de coloração acinzentada e textura arenosa (RODRIGUES et al., 2005; SANTOS et al., 2018).

Na região de Tailândia, o clima é do tipo Am, tropical monçônico, com precipitação do mês mais seco inferior a 60 mm, enquanto as demais regiões Tomé-Açu, Acará e Moju apresentam clima Af, tropical sem estação seca com precipitação do mês mais seco superior a 60 mm. Ambos apresentam temperatura média do mês mais frio do ano acima de 18 °C (ALVARES et al., 2014).

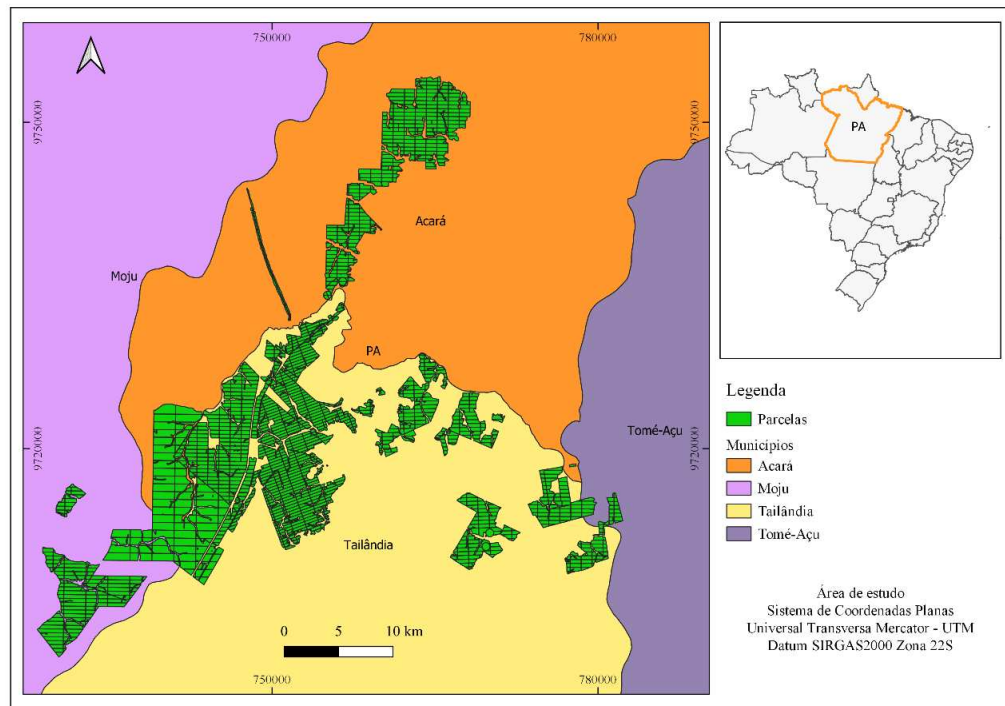


Figura 1 - Localização da área de estudo.

3.2.2 Caracterização dos plantios e organização da base de dados

Os plantios da palma-de-óleo estão estratificados por ano de plantio (1985 a 2018), distribuídos pelas fazendas Crai, Agropalma, Agropar, Amapalma, Castanheira, Trevo/Galileia, Palmares e Roda de Fogo, nas quais foram estabelecidas parcelas de plantios da palmeira (ou talhões) com tamanhos variando entre dois e 118 hectares. Os plantios foram realizados com diferentes materiais genéticos, predominando a variedade Deli X Lamé, e em diferentes espaçamentos entre plantas, com densidade padrão de 143 plantas por hectare.

O inventário florestal é realizado mensalmente, sendo amostradas de 25 a 40 plantas por parcela e quantificados inflorescências em antese (antes da fecundação) e fecunda (já fecundadas), cachos verdes e cachos maduros.

Existem quatro estações meteorológicas instaladas nas fazendas Castanheira, Palmares, Agropalma e Amapalma para coleta dos dados climáticos mensais de temperatura, precipitação, insolação e umidade relativa do ar. Além destes é calculado o déficit hídrico mensal.

Para construção da base de dados, a empresa florestal disponibilizou dados históricos mensais de 2014 a 2022, referentes ao inventário florestal, à produção total de cachos por parcela, aos registros das estações meteorológicas e ao cadastro geral dos plantios. Estes foram relacionados e consistidos, etapa na qual foram corrigidos erros de digitação, excluídas linhas sem valores para todos os atributos, retirados valores discrepantes da produção de cachos, do número de inflorescências em antese e fecunda, e de cachos verdes e maduros. De posse da base consistida, foram pré-definidas as variáveis qualitativas e quantitativas a serem utilizadas na modelagem da produção dos cachos da palma-de-óleo.

3.2.3 Predição da produção de cachos da palma-de-óleo

A predição consistiu em estimar a produção mensal corrente de cachos da palma-de-óleo. Para isto, a base de dados foi organizada de forma a incluir os valores das variáveis climáticas registrados no mês anterior ao mês de predição. Isto, pelo fato de a variável climática do mês atual ainda não ser contabilizada no momento da amostragem e o comportamento climático anterior ter influência sobre a produção atual dos cachos. Sendo gerada uma base contendo 13852 linhas de dados.

Conforme Tabela 1, foram calculados os valores mínimos, médios, máximos, desvio padrão (DP) e coeficiente de variação (CV) das variáveis quantitativas utilizadas no estudo. As variáveis qualitativas incluem oito fazendas, 27 estratos anos de plantio (1985 a 2016), 19 materiais genéticos e seis classes de solos.

Tabela 1 – Análise descritiva das variáveis quantitativas.

Variáveis quantitativas	Sigla	Unidade	Mín	Méd	Máx	DP	CV (%)
Inflorescências em antese	AntHa	n° ha ⁻¹	0,00	4,32	102,96	7,91	182,90
Inflorescências em fecunda	FecHa	n° ha ⁻¹	0,00	3,42	205,29	7,26	211,91
Cacho Verde	CVHa	n° ha ⁻¹	5,24	566,22	1836,12	324,36	57,29
Cacho Maduro	CMHa	n° ha ⁻¹	1,79	96,63	1024,00	95,92	99,26
Peso médio do cacho	PMCa	kg	2,02	14,81	84,06	5,79	39,11
Porcentagem de amostragem	%am	%	0,50	1,06	14,63	0,77	72,69
Idade	-	anos	4	15,33	33	8,08	52,69

Temperatura média do mês anterior	Tmed1	°C	25,62	26,89	28,19	0,70	2,59
Precipitação do mês anterior	PP1	mm	0,40	194,76	670,61	173,36	89,01
Déficit hídrico do mês anterior	Def1	mm	0,00	41,29	168,16	50,70	122,78
Umidade relativa do mês anterior	Ur1	%	67,91	84,80	92,86	5,74	6,77
Insolação do mês anterior	Insol1	W m ⁻²	116,54	157,59	205,44	19,35	12,28
Produção mensal de cachos	ProdHa	t ha ⁻¹	1,00	2,26	21,57	1,44	63,66

As variáveis selecionadas para entrada na primeira camada da RNA incluíram aquelas relacionadas ao cadastro geral das parcelas (Fazenda, AnoPlantio, MaterialGenetico e Solo), inventário do mês atual (AntHa, FecHa, CVHa, CMHa, PMCa), condições climáticas e déficit hídrico referentes ao mês anterior (Tmed1, PP1, Def1, Ur1, Insol1), mês de colheita, idade e porcentagem de amostragem. Como variável de saída foi determinada a produção mensal corrente de cachos por hectare (ProdHa) da palma-de-óleo em nível de parcela.

O software Neuroforest 4.0 (BINOTI et al., 2014) foi utilizado para treinamento e aplicação das redes neuronais artificiais. Estas foram configuradas conforme padrão já definido pelo software, oito neurônios na camada oculta, função de ativação sigmoideal, algoritmo de treinamento resilient propagation, critério de parada 3000 ciclos e separação de 70% dos dados para treinamento e 30% para validação.

De posse das estimativas da produção corrente de cachos por hectare em nível de parcela, calculou-se a produção total de cachos (toneladas) estimada e realizada sob diferentes níveis: toneladas totais de cachos por parcela, ano e respectivo mês de colheita (nível 1), ano e mês de colheita (nível 2), mês de colheita (nível 3) e ano de colheita (nível 4) sem estratificar por parcela, ou seja, para os dados gerais. Diante disso, foram gerados a correlação entre valores observados e estimados e o erro relativo percentual médio para cada nível.

3.2.4 Avaliação das redes neuronais artificiais

Após o processamento, calculou-se as estatísticas correlação linear entre valores preditos e observados (1), bias (2), raiz quadrada do erro quadrático médio percentual (3), erro relativo percentual médio(4), bias percentual (5) e variância (6) para seleção da melhor RNA (ISLAM et al., 2009; CAMPOS; LEITE, 2017). Foram gerados gráficos representando a dispersão dos erros relativos percentuais, a relação entre valores estimados e observados e histograma da frequência dos erros relativos percentuais (4).

$$r_{y\hat{y}} = \frac{\sum_{i=1}^n (Y_i - \bar{Y}) \cdot (\hat{Y}_i - \bar{\hat{Y}})}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2 \sum_{i=1}^n (\hat{Y}_i - \bar{\hat{Y}})^2}} \quad (1)$$

$$bias = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i) \quad (2)$$

$$RQEM\% = \frac{\sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}}}{\bar{Y}} \cdot 100 \quad (3)$$

$$\overline{ER}\% = \frac{1}{n} \sum_{i=1}^n \frac{(\hat{Y}_i - Y_i) \cdot 100}{Y_i} \quad (4)$$

$$bias\% = \frac{bias}{\bar{Y}} \cdot 100 \quad (5)$$

$$Var = \frac{\sum (bias - (Y_i - \hat{Y}_i))^2}{n - 1} \quad (6)$$

Em que: Y_i é o valor observado, \hat{Y}_i é o valor estimado, \bar{Y} é a média dos valores observados, $\bar{\hat{Y}}$ é a média dos valores estimados e n é o número total de observações.

3.2.5 Determinação das variáveis de maior contribuição na modelagem

3.2.5.1 Matriz de correlação

Na modelagem foram utilizadas todas as variáveis do inventário e climáticas existentes na base de dados. No entanto, é importante entender o grau de relação entre essas variáveis e a variável de saída, por isso foi gerada uma matriz de correlação entre as mesmas.

A correlação pode variar de -1 a 1, quanto mais próximo de 1, em módulo, maior a relação entre as variáveis, que pode ser positiva ou negativa. Se as variáveis tendem a aumentar

ou diminuir juntas a correlação é positiva, caso a variável aumente enquanto a outra diminui a correlação é negativa.

O procedimento foi realizado no software R, aplicando o pacote psych (REVELLE, 2022), o qual gera uma matriz de correlação composta por uma diagonal contendo histogramas de distribuição de cada variável, na parte superior da diagonal as correlações e na porção inferior gráficos representando a relação entre as variáveis.

3.2.5.2 Aplicação do algoritmo de Garson

O algoritmo de Garson (1991) gera a importância relativa de cada variável de entrada em relação ao valor predito na modelagem. Para tanto, por meio da aplicação de redes neurais artificiais, o algoritmo utiliza os valores absolutos dos pesos entre os neurônios da camada de entrada e oculta e entre a camada oculta e de saída para quantificar a importância de cada variável (OLDEN; JACKSON, 2002; FREITAS et al., 2020).

Com o intuito de definir as variáveis de maior contribuição na modelagem, foi treinada uma rede neuronal de mesma configuração à utilizada anteriormente neste estudo, adotando todas as variáveis qualitativas e quantitativas disponíveis, e, posteriormente, aplicado o algoritmo Garson que utilizou os pesos gerados nesse treino. O processamento foi realizado no software R, empregando o pacote Neuralnet (FRITSCH et al., 2019).

3.2.5.3 Testes adicionais de modelagem

Após definição das variáveis mais relacionadas com a produção de cachos por hectare, foram feitos dois novos testes de modelagem, retirando algumas variáveis do inventário e elementos climáticos. Estes foram comparados com a modelagem original, calculando as estatísticas correlação entre valores observados e estimados e o erro relativo percentual médio.

3.3 Resultados e discussão

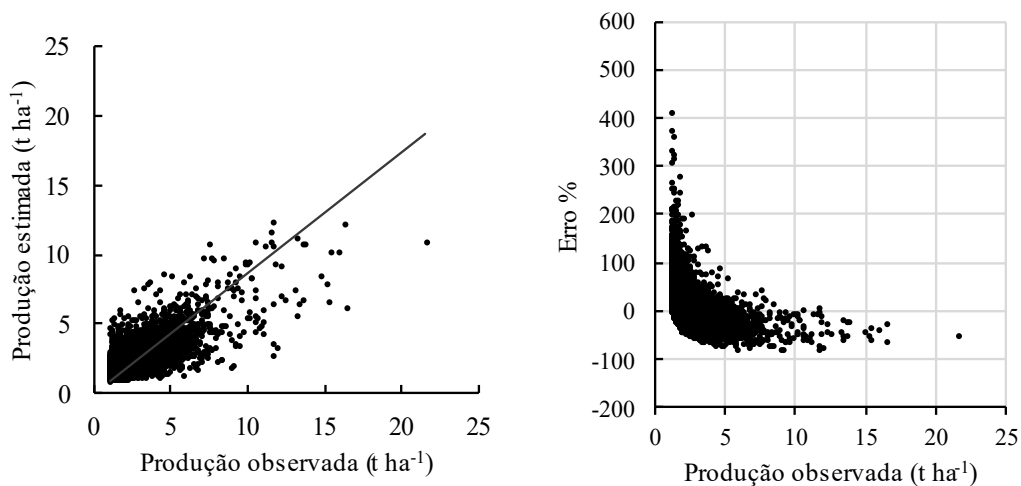
3.3.1 Predição da produção de cachos

A RNA selecionada apresentou estatísticas de erro semelhantes no treino e validação, com erro relativo percentual médio de aproximadamente 13%, RQEM% na faixa dos 40%, bias de aproximadamente zero e variância em torno de um, representando a variabilidade das predições. Quando se avalia a correlação entre valores observados e estimados há uma diferença, sendo acima de 0,7 para o treino e de 0,6 para a validação (Tabela 2).

Tabela 2 – Estatísticas calculadas para avaliação da acurácia da RNA nas fases de treino e validação.

Tipo	Bias	Bias %	RQEM%	Var	$\overline{ER\%}$	$r_{y\hat{y}}$
Treino	0,000001	0,000060	42,29	0,92	12,62	0,74
Validação	-0,009437	-0,420264	47,86	1,15	13,36	0,67

Os gráficos referentes a dispersão do erro relativo percentual mostram que, a maioria dos erros varia entre -100 e 100%, principalmente nas produções entre 1 e 5 toneladas por hectare. Nas menores produções alguns pontos começam a se distanciar da tendência geral, configurando erros acima de 100% (Figura 2).



A

B

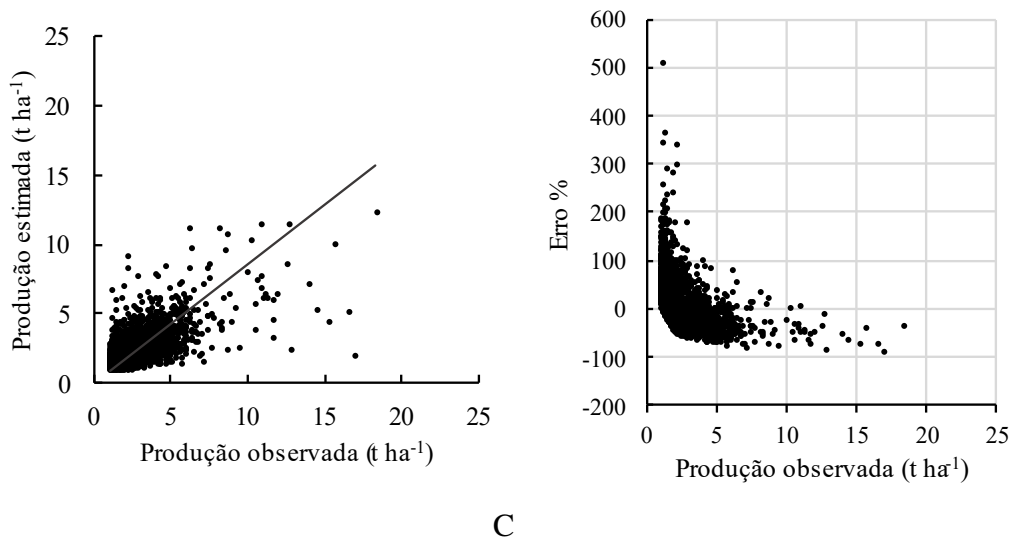


Figura 2 – Relações entre produções estimadas e observadas e entre erros relativos percentuais e produções observadas para treino (A, B) e validação (C, D) da RNA.

A distribuição dos erros relativos percentuais é melhor visualizada quando agrupados em classes (Figura 3). Assim, foram definidas classes variando de -100 a 100%, somando pouco mais de 80% dos erros entre -40 e 60%, tanto no treino quanto na validação (Figura 3).

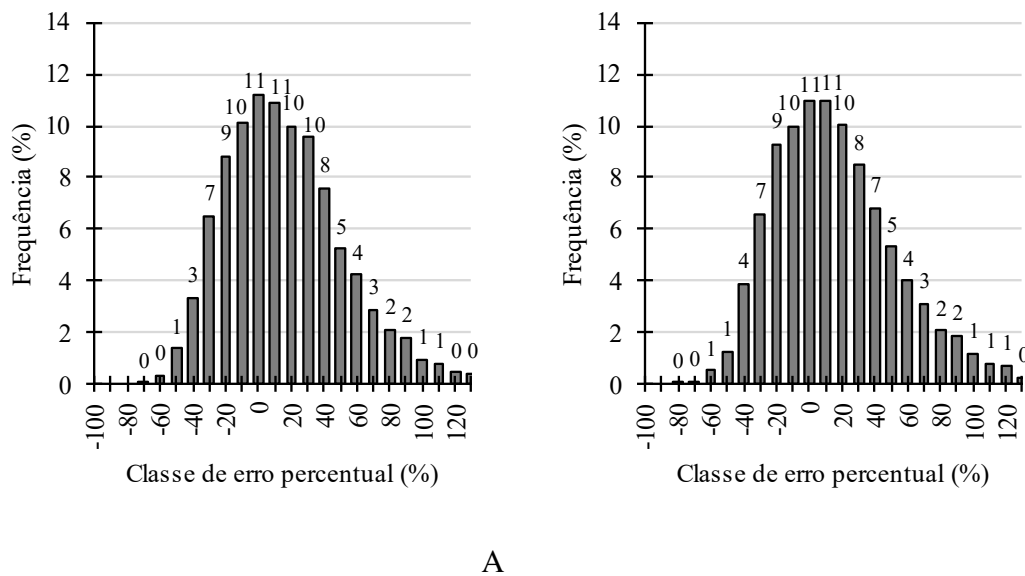
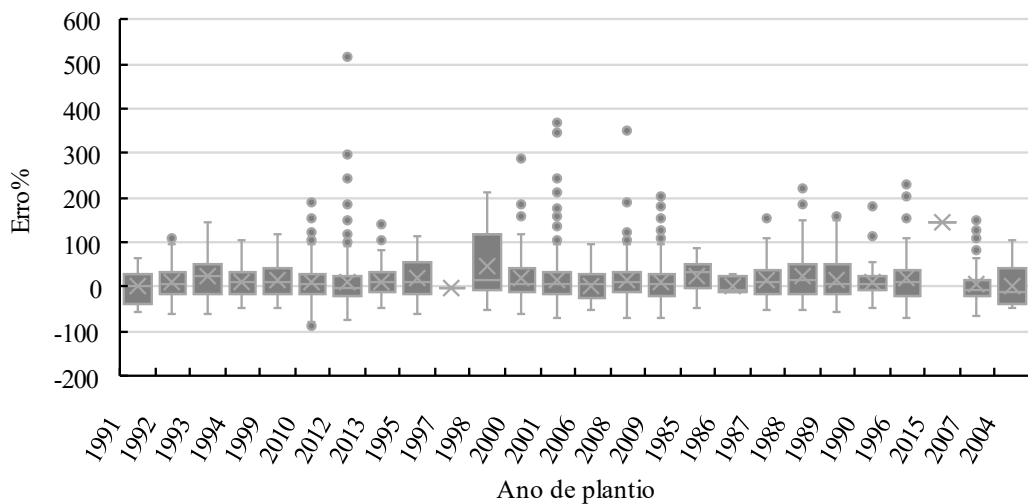


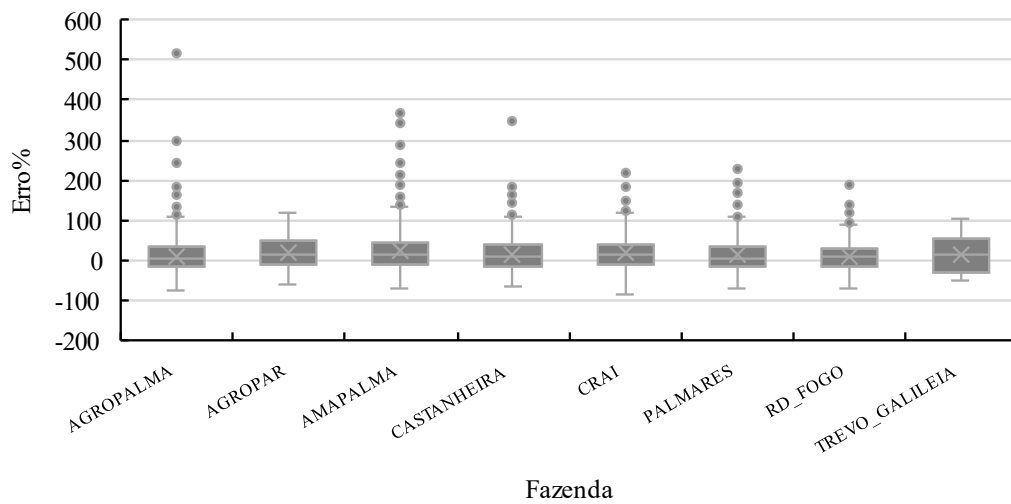
Figura 3 – Percentual de casos por classe de erro relativo percentual das estimativas de produções de cachos obtidos com os dados de treino (A) e validação (B).

O comportamento dos erros percentuais em função das variáveis qualitativas revela a presença de outliers na maioria das categorias de cada variável (Figura 4). Os estratos anos de plantio 2001, 2009, 2010 e 2012, caracterizados pelas maiores áreas de plantio, apresentaram grande parte dos outliers, ou seja, erros acima de 100% (Figura 4A). O mesmo padrão foi encontrado para as fazendas, aquelas de maior área geraram outliers, ao contrário das fazendas

de áreas pequenas, Agropar e Trevo/Galileia (Figura 4B). Os materiais genéticos Deli Lamé e Deli Yangambi, presentes na maioria dos plantios, compreenderam erros discrepantes da tendência geral (-100 a 100%). A variedade Deli Dami, de pouca representatividade na base de dados, apresentou todos os erros entre 80 e 200% (Figura 4C). Quanto ao solo, observa-se que o Argissolo (pad1, pad2, pad3, pad4, pad5), predominante na região, concentra os erros acima de 100%. Tanto no material genético quanto no solo, não existe um padrão de comportamento do erro, o que implica na influência de outros fatores nessas variáveis.



A



B

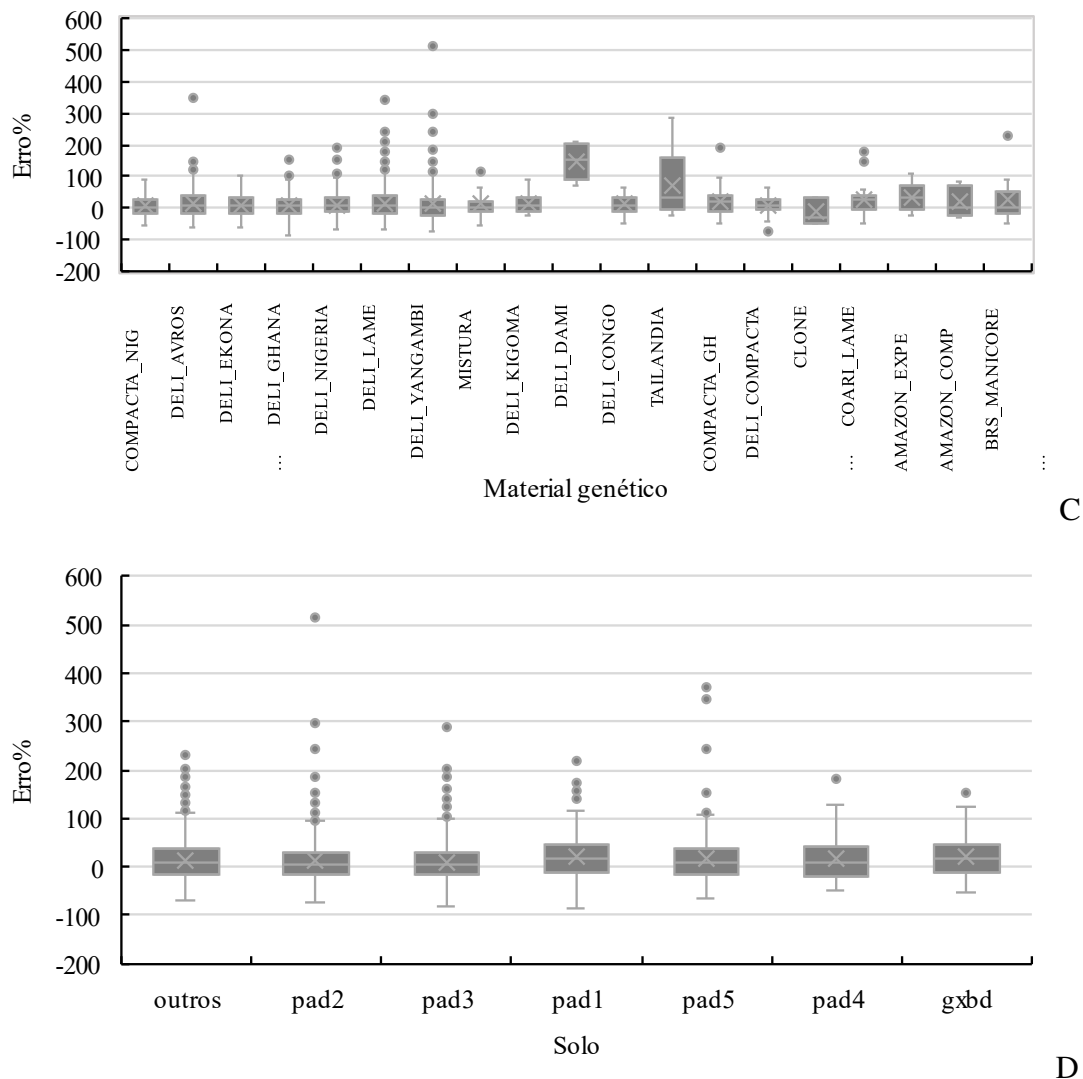
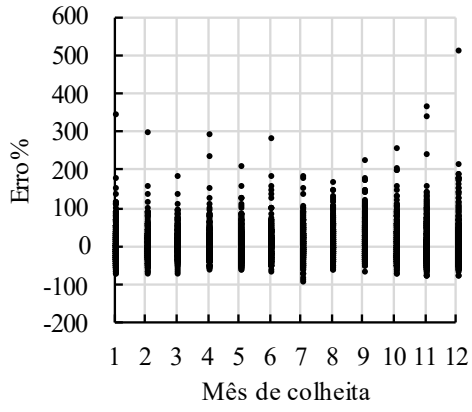
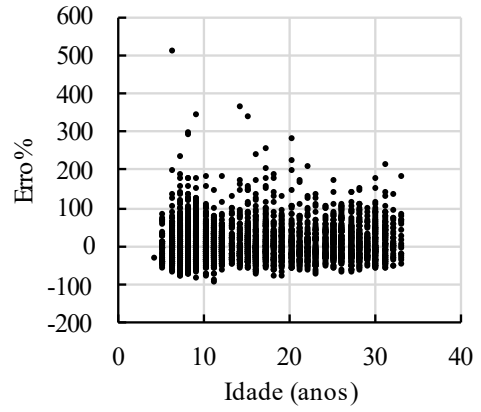


Figura 4 - Dispersão dos erros relativos percentuais avaliados por variável qualitativa na fase de validação da RNA.

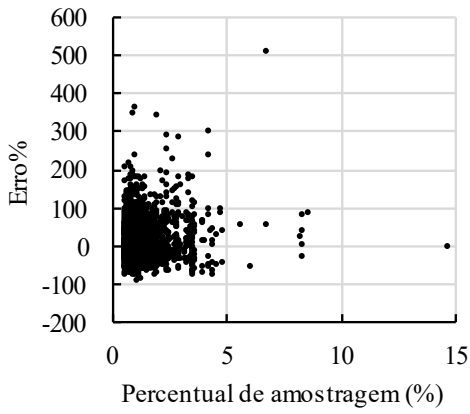
Quando se avalia a dispersão residual percentual em função das variáveis quantitativas, o percentual de amostragem merece destaque, visto os maiores erros nas menores porcentagens, especificamente entre 0 e 1% (Figura 5C). Variáveis do inventário como número de antes, fecundas e cachos maduros (Figura 5E, 5F e 5H) seguem o mesmo padrão de comportamento, nos menores valores a RNA apresenta maior erro. Ao realizar o inventário mensal, em muitos casos, ocorre identificação de um número muito pequeno de cachos e inflorescências, ou até mesmo, não são verificadas sua presença. O que não implica em produção nula naquele mês, já que a produção é quantificada em um determinado momento. Esses valores mínimos ou nulos não representam ou estão relacionados diretamente com as toneladas reais de cachos que chegam à fábrica no mês corrente. Assim, ao relacionar estas variáveis com a produção de cachos realizada para o mês corrente, a RNA tende a errar mais, pois não consegue encontrar uma relação evidente. Quanto às variáveis climáticas, não há uma tendência clara.



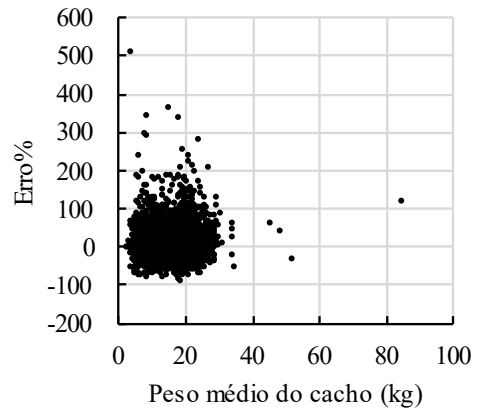
A



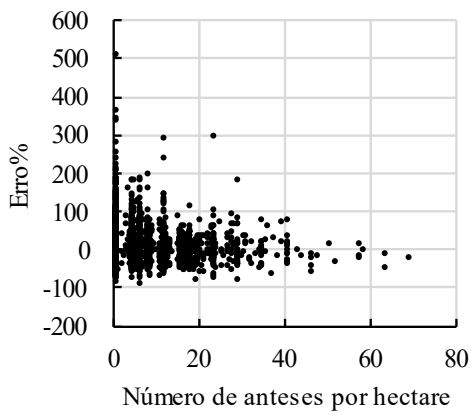
B



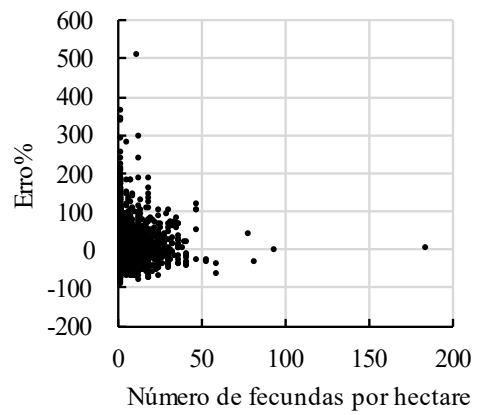
C



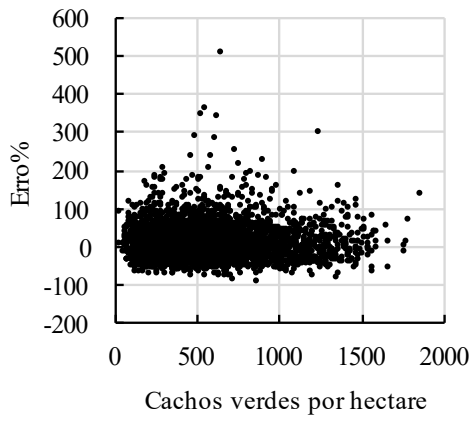
D



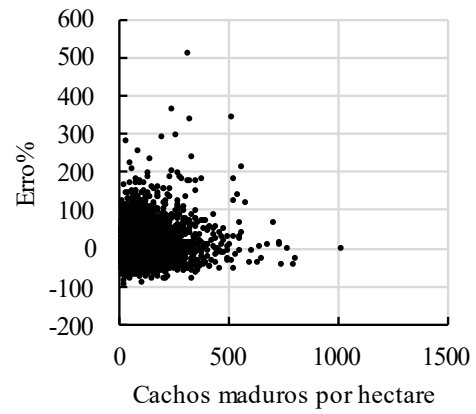
E



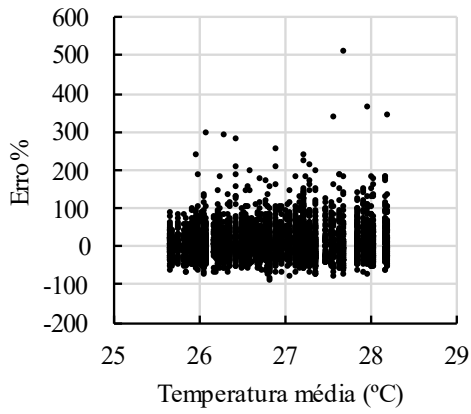
F



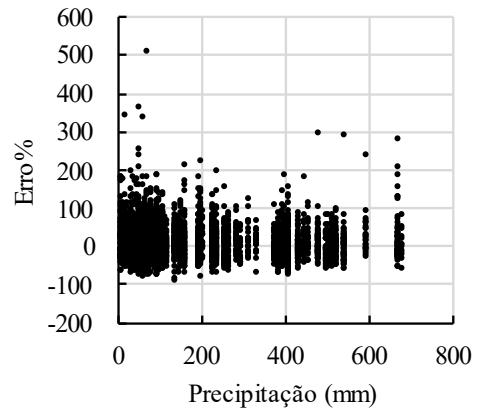
G



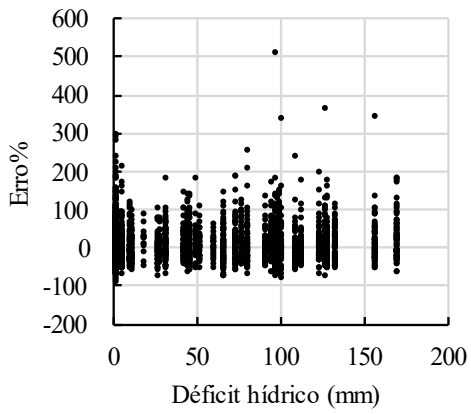
H



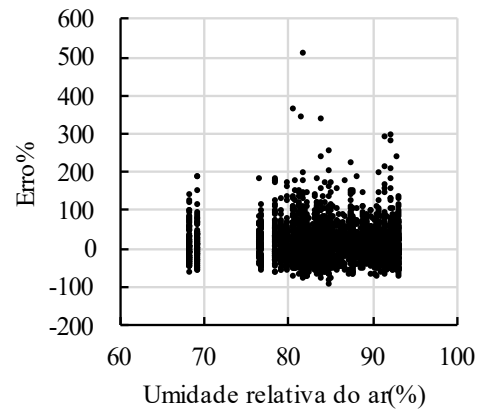
I



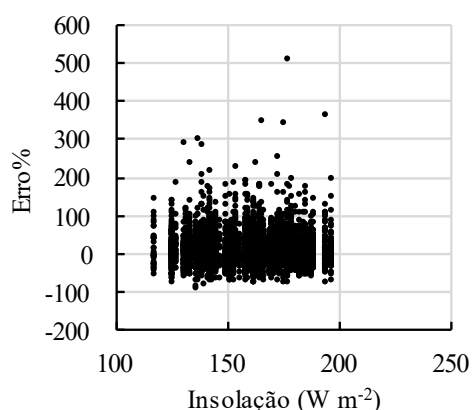
J



K



L



M

Figura 5 – Dispersão dos erros relativos percentuais avaliados por variável quantitativa na fase de validação da RNA.

A modelagem em nível de parcela sofre influência do banco de dados de amostragem, procedimento operacional no qual a parcela é objeto do manejo, constituindo o principal fator de interferência na precisão das estimativas (KHAN et al., 2021; MOHD NAIN et al., 2022), por apresentar alta variabilidade, o que é comprovado ao analisar as variáveis antese, fecunda, cacho verde, cacho maduro, porcentagem de amostragem, assim como os valores da produção de cachos ao longo dos anos.

O processo de coleta dos dados de inventário é um dos grandes desafios do campo, devido à dificuldade em visualizar cachos e inflorescências, especialmente, quando se trata das plantas maiores, ocasionando erros na contagem. Isto gera menor confiabilidade nos dados que chegam do campo ao escritório para serem consistidos e processados. Tal questão foi discutida em estudo de Chapman et al. (2018), o qual associou a menor precisão das estimativas de rendimento dos cachos de frutos frescos a erros de registros no banco de dados. A redução do impacto dos erros de amostragem na modelagem depende de alterações no planejamento dos procedimentos operacionais, como, por exemplo, a aplicação de porcentagens maiores de amostragem nas áreas de elevada heterogeneidade (SOARES et al., 2011). Assim, a RNA será capaz de compreender melhor a variabilidade produtiva e os fatores que a influenciam.

Avaliando o desempenho das estimativas sob diferentes níveis, ocorre um aumento da correlação e diminuição dos erros percentuais conforme as estimativas são avaliadas de forma geral, havendo uma alta assertividade da RNA ao estimar a produção total em cada mês (0,99 e 1,9%), bem como em cada ano de colheita (0,99, 1,5%). Considerando que um dos objetivos

é obter estimativas da produção mensal de cachos com maior confiabilidade, a análise para o nível 3, comprova a eficácia da modelagem utilizando redes neurais artificiais (Tabela 3).

Tabela 3 – Análise da produção total de cachos estimada por ano e mês de colheita

Produção total cachos (ton)	Nível	Estrato	$r_{\hat{y}y}$	$\overline{ER\%}$
Ano e mês de colheita	1	Parcela	0,75228	13,3%
Ano e mês de colheita	2	Geral	0,99103	6,3%
Mês de colheita	3	Geral	0,99725	1,9%
Ano de colheita	4	Geral	0,99615	1,5%

3.3.2 Variáveis de maior contribuição na modelagem

Adotando como critério correlações acima de 0.2 (módulo), a produção de cachos por hectare (ProdHa) tem maior relação com as variáveis mês de colheita (Mês), porcentagem de amostragem (%am), cachos maduros (CMHa), temperatura média (Tmed1), déficit hídrico (Def1), precipitação (PP1) e umidade relativa (Ur1) (Figura 6).

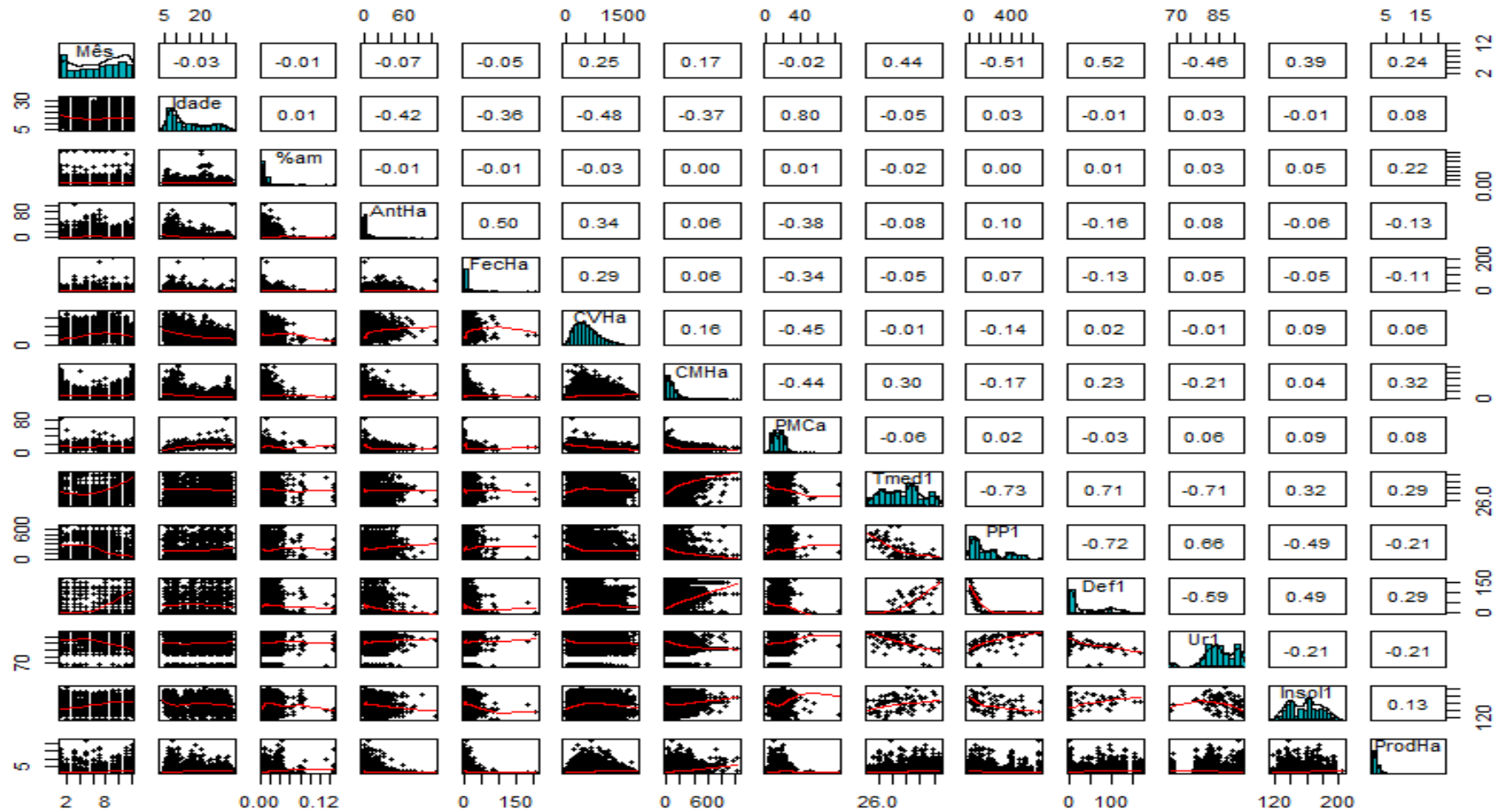


Figura 6 – Correlação de Pearson entre as variáveis quantitativas utilizadas na modelagem.

De acordo com o algoritmo de Garson as variáveis de maior contribuição na modelagem são o ano de plantio, material genético, solo, fazenda (qualitativas), inflorescências antese e fecunda, cachos maduros, peso médio do cacho e porcentagem de amostragem (quantitativas).

De acordo com a Tabela 4, que resume as variáveis (marcadas em X) utilizadas em cada teste adicional de modelagem, e suas respectivas estatísticas correlação e erro relativo médio percentual, é perceptível que os dois novos testes (Teste 1 e Teste 2) não apresentaram resultados discrepantes do teste original (Teste 0).

Tabela 4 – Variáveis de entrada utilizadas em cada teste e respectivas estatísticas da modelagem (correlação, erro relativo percentual médio) para treino (T) e validação (V).

Variáveis		Teste 0	Teste 1	Teste 2
AnoPlantio		X	X	X
Fazenda		X	X	X
MaterialGenetico		X	X	X
Solo		X	X	X
Mês		X	X	X
Idade		X		
%am		X	X	X
AntHa		X		
FecHa		X		
CVHa		X		
CMHa		X	X	X
PMCa		X		
Tmed1		X	X	
PP1		X		
Def1		X	X	
Ur1		X		
Insol1		X	X	
$r_{y\hat{y}}$	T	0,74	0,71	0,71
	V	0,67	0,67	0,62
$\overline{ER\%}$	T	12,62	13,38	13,68
	V	13,36	14,66	15,92

Tanto o algoritmo Garson como os testes realizados mostraram que as variáveis climáticas não exerceram tanta influência sobre as estimativas. O teste 2 que excluiu todas as variáveis climáticas obteve valores estatísticos semelhantes aos outros testes (Tabela 4). Este resultado pode indicar que a variação climática entre os meses não foi expressiva o suficiente para afetar o aprendizado da RNA ou uma menor qualidade da base de dados climática. Isto porque, as estações meteorológicas apresentaram problemas técnicos em determinados períodos, diminuindo a confiabilidade das variáveis coletadas. Por isso, é importante investir em equipamentos e técnicas que incrementem a qualidade dos dados climáticos disponibilizados (HOFFMANN et al., 2014, 2017; KARTIKA et al., 2016). Aliado a isto, o uso do fator climático apenas do mês anterior ao mês de predição não foi suficiente para expressar a interferência do clima na produção de cachos da palma-de-óleo, visto que o desenvolvimento das inflorescências ocorre 32 meses após a germinação das sementes, e nesse período a influência do elemento climático é cumulativa (ADAM et al., 2005; FARIAS, 2014). Sendo interessante em estudos futuros a aplicação das variáveis climáticas defasadas e acumuladas em até 32 meses anteriores ao mês de predição.

A manutenção da coleta e uso das variáveis climáticas é imprescindível para estudos de modelagem da palma-de-óleo, visto a influência do clima sobre o desenvolvimento da espécie. Em casos de eventos aleatórios, a inclusão de variáveis como temperatura, radiação solar, precipitação ou déficit hídrico torna-se indispensável. O déficit hídrico, por exemplo, é o principal fator limitante da produção dos cachos, isto porque, retarda a iniciação floral, aumenta a taxa de aborto e, conseqüentemente, diminui a produção de frutos (BOARI et al., 2010; CORLEY; TINKER, 2016; PIRKER et al., 2016; RHEBERGEN et al., 2016; MURUGESAN et al., 2017; MONZON et al., 2022). Esse mesmo efeito é visualizado em períodos de altas precipitações, havendo uma correlação negativa com a produção de cachos de frutos frescos e a produção de óleo no cacho, já que o excesso de chuvas durante o desenvolvimento das inflorescências promove o abortamento das flores e, conseqüente, menor produção de cachos e de óleo, com efeitos na taxa de extração industrial (TEI) (SILVA, 2006).

A palma-de-óleo é uma espécie exigente de condições climáticas específicas, que ao não serem atendidas ocasiona baixos rendimentos anuais. A região da Amazônia, que abrange o presente estudo, caracteriza-se por médias anuais de temperatura entre 24°C e 28°C, deficiência hídrica inferior a 200 mm, umidade relativa do ar entre 75% e 90%, precipitações superiores a 100 mm e altos níveis de radiação solar (1600 horas/ano) (BOARI et al., 2010), tais condições favorecem o estabelecimento da palmeira na região. No entanto, as constantes

mudanças climáticas tendem a alterar as condições normais e, conseqüentemente, os rendimentos da cultura.

O uso dos estratos anos de plantio e fazenda na modelagem, fornece informações específicas do gerenciamento dos plantios, além de refletir as mudanças ambientais, por serem afetados diretamente. Tais estratos expressam, em sua maioria, um padrão de picos de produção ao longo do ano de colheita, que acontece nos meses de janeiro, novembro e dezembro e coincidem com os picos de temperatura média e insolação, que constituem um dos elementos climáticos de maior influência sobre a produtividade do dendê, juntamente com a precipitação pluviométrica (HOFFMANN et al., 2014; CORLEY; TINKER, 2016; MONZON et al., 2022). Monzon et al. (2022) avaliaram a variação da produtividade do dendê no espaço (estratos de plantio) e no tempo (ano e mês) na região da Indonésia. Em seu estudo, os picos de rendimento de cachos mensais diferiram nas quatro regiões avaliadas, ocorrendo no início (janeiro), meio (maio – agosto) ou final do ano (novembro). Tais rendimentos foram intimamente associados ao estresse hídrico. Logo, percebe-se que a produção mensal de cachos da palma-de-óleo é dependente de aspectos próprios da região de implantação da espécie, bem como do comportamento climático do local.

A importância das variáveis categóricas tipo de solo e material genético corrobora com pesquisas realizadas fora do Brasil, como discutido em estudo de revisão de Mohd Nain et al. (2022), os quais dividem as variáveis em três categorias: ambiente, fenótipo e genótipo; e destacam o genótipo como um fator amplamente utilizado nos artigos analisados, e o solo e fertilização como o segundo fator ambiental mais abordado nos estudos, perdendo apenas para o clima.

Dentre as variáveis do inventário, a identificação e contabilização de cachos maduros, é essencial para a modelagem, visto sua relação direta com o rendimento do óleo de palma (MONZON et al., 2022). Após a colheita dos cachos, estes devem ser beneficiados e processados o mais rápido possível, para evitar acidez elevada no óleo extraído e, conseqüentemente, perda de qualidade do produto final (MULLER et al., 2006; TELES, 2014; CORLEY; TINKER, 2016). Normalmente, a classificação dos cachos de frutos frescos é feita manualmente, de acordo com o julgamento subjetivo do responsável, e é influenciada pela altura das palmeiras, posição do fruto e direção da luz solar. Além disso é observado o número de frutos destacados no chão para confirmar a maturação do cacho. Este método não é tão eficiente, já que não consegue capturar a qualidade interna do fruto, apresentando alta taxa de erros, que afetam diretamente a predição de rendimento do óleo de palma (ALFATNI et al.,

2018, 2020; ISHAK et al., 2019; MAKKY; CHERIE, 2021). Por isso, o uso de métodos alternativos e precisos de classificação de cachos maduros é importante para estimar toneladas de cachos por área plantada, como, por exemplo, classificação via imagem.

3.4 Conclusão

Redes neuronais artificiais clássicas, com uma única camada de neurônios oculta e configuração simples são capazes de prever a produção corrente de cachos da palma-de-óleo, apresentando desempenho satisfatório para fins de inventário. Assim, constitui uma técnica computacional passível de aplicação em estudos de modelagem da palma-de-óleo utilizando fatores ambientais como variáveis preditoras.

3.5 Referências bibliográficas

ADAM, H.; JOUANNIC, S.; ESCOUTE, J.; et al. Reproductive developmental complexity in the African oil palm (*Elaeis guineensis*, Arecaceae). **American Journal of Botany**, v. 92, n. 11, p. 1836–1852, 2005.

ADIZUE, U. L.; NWANYA, S. C.; OZOR, P. A. Artificial neural network application to a process time planning problem for palm oil production. **Engineering and Applied Science Research**, v. 47, n. 2, p. 161–169, 2020.

AGROPALMA. **Relatório de sustentabilidade**. Helikonia, 2019.

ALFATNI, M. S. M.; MOHAMED SHARIFF, A. R.; BEJO, S. K.; SAAED, O. M. BEN; MUSTAPHA, A. Real-time oil palm FFB ripeness grading system based on ANN, KNN and SVM classifiers. **IOP Conference Series: Earth and Environmental Science**, v. 169, n. 1, 2018.

ALFATNI, M. S. M.; MOHAMED SHARIFF, A. R.; SAAED, O. M. BEN; ALBHBAH, A. M.; MUSTAPHA, A. Colour Feature Extraction Techniques for Real Time System of Oil Palm

Fresh Fruit Bunch Maturity Grading. **IOP Conference Series: Earth and Environmental Science**, v. 540, n. 1, 2020.

ALVARES, C. A.; STAPE, L.; SENTELHAS, P. C.; GONÇALVES, L. D. M.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 2014.

BINOTI, D.; SILVA; LEITE, H. G.; LOPES, P. F.; LOPES, S. S. P., NeuroForest, Versão 4.0. Viçosa: DAP Florestal, 2014. (Software). www.neuroforest.ucoz.com

BINOTI, M. L. M. DA S. **Redes neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto**, 2010. Dissertação (Mestrado), Viçosa: Universidade Federal de Viçosa.

BINOTI, M. L. M. DA S.; LEITE, H. G.; BINOTI, D. H. B.; GLERIANI, J. M. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **Cerne**, v. 21, n. 1, p. 97–105, 2015. Federal University of Lavras.

BOARI, A. J.; BARROS, A. H.; GONÇALVES, A. O.; et al. **Zoneamento agroecológico, produção e manejo para a cultura da Palma-de-óleo na Amazônia**. 1st ed. Rio de Janeiro: Agência 2A comunicação, 2010.

BRIASSOULIS, H. **The Web Book of Regional Science Sponsored by Analysis of Land Use Change: Theoretical and modeling approaches by Helen Briassoulis**. 2nd ed. WVU Research Repository, 2020.

CAMPOS, J.C.C.; LEITE, H.G. Mensuração Florestal: Perguntas e Respostas. 5ª. Ed.; Viçosa: Editora UFV, 2017. 636 p.

CHAPMAN, R.; COOK, S.; DONOUGH, C.; et al. Using Bayesian networks to predict future yield functions with data from commercial oil palm plantations: A proof of concept analysis. **Computers and Electronics in Agriculture**, v. 151, n. October 2017, p. 338–348, 2018.

CORLEY, R. H. V; TINKER, P. B. **The Oil Palm**. BlackWell Science Ltd, 2016.

DROST, S.; RIJK, G.; PIOTROWSKI, M. Palm Oil Production in Brazilian Amazon Threatens NDPE Commitments. **Chain Reaction Research**, 2021. Disponível em: www.chainreactionresearch.com.

FARIAS, M. P. **Caracterização do desenvolvimento floral de *Elaeis oleifera* (H. B. K.) Cortés**, 2014. Dissertação (Mestrado), Lavras: Universidade Federal de Lavras – UFLA.

FAS - FOREIGN AGRICULTURAL SERVICES. Oil, Palm explorer. Disponível em: <<https://ipad.fas.usda.gov/cropexplorer/cropview/commodityView.aspx?cropid=4243000>>. Acesso em: 9/3/2022.

FREITAS, E. C. S. DE; PAIVA, H. N. DE; NEVES, J. C. L.; MARCATTI, G. E.; LEITE, H. G. Modeling of eucalyptus productivity with artificial neural networks. **Industrial Crops and Products**, v. 146, 2020.

FRITSCH, S.; GUENTHER, F. WRIGHT, M. N. (2019). neuralnet: Training of Neural Networks. R package version 1.44.2. <https://CRAN.R-project.org/package=neuralnet>

GARSON, G.D. Interpreting neural network connection weights. **Artificial Intelligence Expert**, v. 6, n.4, p. 46-51, 1991.

GORGENS, E. B.; LEITE, H. G.; SANTOS, H. DO N.; GLERIANI, J. M. Estimação do volume de árvores utilizando redes neurais artificiais. **Revista Árvore**, v. 33, n. 6, p. 1141–1147, 2009.

HOFFMANN, M. P.; DONOUGH, C. R.; COOK, S. E.; et al. Yield gap analysis in oil palm: Framework development and application in commercial operations in Southeast Asia. **Agricultural Systems**, v. 151, p. 12–19, 2017. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.agsy.2016.11.005>>.

HOFFMANN, M. P.; VERA, A. C.; WIJK, M. T. VAN; et al. Simulating potential growth and yield of oil palm (*Elaeis guineensis*) with PALMSIM: Model description, evaluation and application. **Agricultural Systems**, v. 131, p. 1–10, 2014. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.agsy.2014.07.006>>.

IGUODALA, E. A.; THOMAS, E. J.; EDOKPAYI, A; TITILOLA, O. A Mixed model Analysis of a fertilizer experiment on oil palm in Nigeria. **Agricultural Sciences**, v. 7, n. 8, p. 521 – 530, 2016.

ISHAK, H.; SHIDDIQ, M.; FITRA, R. H.; YASMIN, N. Z. Ripeness Level Classification of Oil Palm Fresh Fruit Bunch Using Laser Induced Fluorescence Imaging. **Journal of Aceh Physics Society**, v. 8, n. 3, p. 84–89, 2019.

ISLAM, M. N.; KURTTILA, M.; MEHTÄTALO, L.; HAARA, A. Analyzing the effects of inventory errors on holding-level forest plans: The Case of measurement error in the basal area of the dominated tree species. **Silva Fennica**, v. 43, n. 1, p. 71–85, 2009.

KARTIKA, N. D.; ASTIKA, I. W.; SANTOSA, E. Oil palm yield forecasting based on weather variables using artificial neural network. **Indonesian Journal of Electrical Engineering and Computer Science**, v. 3, n. 3, p. 626–633, 2016.

KHAN, N.; KAMARUDDIN, M. A.; SHEIKH, U. U.; YUSUP, Y.; BAKHT, M. P. Oil palm and machine learning: Reviewing one decade of ideas, innovations, applications, and gaps. **Agriculture (Switzerland)**, v. 11, n. 9, p. 1–26, 2021.

LEITE, H. G.; ALCÂNTARA, A. E. M.; BINOTI, D. H. B.; OLIVEIRA NETO, R. R.; SILVA, M. L. M. Comparação entre predição e projeção da distribuição de diâmetros de povoamentos de Eucalipto submetidos a desbastes. **Revista Árvore**, v. 37, n. 2, p. 321–328, 2013.

LOUREIRO, G. H. **Redes neurais artificiais e modelos de regressão para estimação de variáveis dendrométricas em *Pinus taeda* L: Análise comparativa.**, 2016 Tese (Doutorado), Curitiba: Universidade Federal do Paraná.

MAKKY, M.; CHERIE, D. Pre-harvest oil palm FFB nondestructive evaluation technique using thermal-imaging device. **IOP Conference Series: Earth and Environmental Science**, v. 757, n. 1, 2021.

MOHD NAIN, F. N.; AHAMED HASSAIN MALIM, N. H.; ABDULLAH, R.; et al. A Review of an Artificial Intelligence Framework for Identifying the Most Effective Palm Oil Prediction. **Algorithms**, v. 15, n. 6, p. 218, 2022.

MONZON, J. P.; JABLOUN, M.; COCK, J.; et al. Influence of weather and endogenous cycles on spatiotemporal yield variation in oil palm. **Agricultural and Forest Meteorology**, v. 314, 2022.

MULLER, A. A.; FURLAN, J. J.; CELESTINO, P. F. **A Embrapa Amazonia Oriental e o Agronegócio do dendê no Pará.** Belém: Embrapa Amazônia Oriental, 2006.

MURUGESAN, P.; ASWATHY, G. M.; SUNIL KUMAR, K.; et al. Oil palm (*Elaeis guineensis*) genetic resources for abiotic stress tolerance: A review. **Indian Journal of Agricultural Sciences**, v. 87, n. 5, p. 571–579, 2017.

OIT - ORGANIZAÇÃO INTERNACIONAL DO TRABALHO. **Cadeia produtiva do óleo de palma - Avanços e desafios rumo à promoção do trabalho decente: Análise Situacional**. 2020.

OLDEN, J. D.; JACKSON, D. A. Illuminating the “black box”: Understanding variable contributions in artificial neural networks. **Ecological Modelling**, v. 154, p. 135–150, 2002.

PIRKER, J.; MOSNIER, A.; KRAXNER, F.; HAVLÍK, P.; OBERSTEINER, M. **What are the limits to oil palm expansion? Global Environmental Change**, v. 40, p. 73–81, 2016. Elsevier Ltd.

REVELLE, W. (2022) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych> Version = 2.2.5.

RHEBERGEN, T.; FAIRHURST, T.; ZINGORE, S.; et al. Climate, soil and land-use based land suitability evaluation for oil palm production in Ghana. **European Journal of Agronomy**, v. 81, p. 1–14, 2016. Elsevier B.V.

RODRIGUES, T. E.; SILVA, J. M. L.; SILVA, B. N. R.; VALENTE, M. A.; GAMA, J. R. N. F.; SANTOS, E. S.; ROLLIM, P. A. M.; PONTE, F. C. **Caracterização e Classificação dos Solos do Município de Tailândia, Estado do Pará**. Belém: Embrapa Amazônia Oriental, 2005.

SANTOS, H. G. DOS; JACOMINE, P. K. T.; ANJOS, L. H. C. DOS; OLIVEIRA, V. A.; LUMBRERAS, J. F.; COELHO, M. R.; ALMEIDA, J. A.; ARAÚJO FILHO, J. C.; OLIVEIRA, J. B.; CUNHA, T. J. F. **Sistema brasileiro de classificação de solos**. Brasília: Embrapa solos, 2018.

SILVA, J. S. O. **Produtividade de óleo de palma na cultura do dendê na Amazônia Oriental: influência do clima e do material genético**, 2006. Dissertação (Mestrado), Viçosa: Universidade Federal de Viçosa - UFV.

SOARES, C.P.B.; PAULA NETO, F.; SOUZA, A.L. **Dendrometria e inventário florestal**. 2ª. Ed. Viçosa: Editora UFV. 2011. 272 p.

TELES, D. A. A. **Características físicas e rendimento mensal em Óleo de cachos de duas cultivares de dendezeiro cultivadas, sob irrigação, no Cerrado do Distrito Federal**., 2014. Dissertação (Mestrado), Brasília: Universidade de Brasília - UNB.

WATSON-HERNÁNDEZ, F.; GÓMEZ-CALDERÓN, N.; SILVA, R. P. DA. Oil Palm Yield Estimation Based on Vegetation and Humidity Indices Generated from Satellite Images and Machine Learning Techniques. **AgriEngineering**, v. 4, n. 1, p. 279–291, 2022.

YOUSEFI D.B., M.; MOHD RAFIE, A. S.; ABD AZIZ, S.; AZRAD, S.; MASRI, M. M. M.; SHAHI, A.; MARZUKI, O. F. Classification of oil palm female inflorescences anthesis stages using machine learning approaches. **Information Processing in Agriculture**, v. 8, n. 4, p. 537–549, 2021.

4 CAPÍTULO 3 – PROJEÇÃO DA PRODUÇÃO DE CACHOS DA PALMA-DE-ÓLEO UTILIZANDO REDE NEURONAL ARTIFICIAL

4.1 Introdução

O óleo de palma ou dendê tem forte participação na economia mundial, muito em função do seu elevado rendimento, sendo o óleo vegetal mais produzido e comercializado no mundo (MBA et al., 2015; DEY et al., 2021). A crescente demanda pelo óleo e seus derivados exige o aumento da produção e aproveitamento ótimo dos recursos associados à produtividade da cultura. Por isso, entender as relações entre as variáveis que influenciam o rendimento do óleo de palma é fundamental para a estimativa da produção futura e definição das ações de manejo que otimizem o processo produtivo (MOHD NAIN et al., 2022).

Prever a produção da espécie não é uma tarefa simples, principalmente, nos últimos anos, visto a dificuldade de controlar fatores ambientais, econômicos e políticos que afetam a produtividade e estabilidade no mercado (BARTHEL et al., 2018). Quanto maior o horizonte de tempo para previsão da produção, menor o controle sob tais variáveis, maiores as chances de equívocos nas estimativas e, conseqüentemente, na tomada de decisão. Sendo importante que o gestor tenha a capacidade de se adaptar estrategicamente à possíveis eventos inesperados (ADIZUE et al., 2020; MOHD NAIN et al., 2022). No caso da palma-de-óleo, espécie exigente de condições ambientais específicas, principalmente climáticas (CORLEY; TINKER, 2016; PIRKER et al., 2016), há maior probabilidade de ocorrência de eventos extraordinários e que afetam a produção futura, como variações extremas de temperatura e precipitação ao longo do ano, o que nos últimos anos vêm sendo uma realidade (SILVA et al., 2019; AVILA-DIAZ et al., 2020).

Normalmente, a tomada de decisão para resolução de problemas extraordinários é feita em tempo curto, de modo empírico, intuitivo, sem estudo prévio e apoio técnico computacional, ocasionando em resultados abaixo do esperado (CHAPMAN et al., 2018). Por isso, utilizar ferramentas computacionais como parte integrante do processo de tomada de decisão é excelente forma de conferir ao gestor maior confiabilidade, permitindo avaliação técnica minuciosa do problema, que, geralmente, envolve uma série de variáveis (SEMBIRING et al., 2017).

Técnicas de inteligência artificial, como as redes neurais artificiais, têm sido aplicadas para resolução de problemas complexos, pela sua capacidade de generalização e menor sensibilidade á ruídos (BAYAT et al., 2020). Além disso, são capazes de trabalhar com grande conjunto de dados, reconhecer padrões de comportamento, e utilizá-los para previsão de respostas (CHAPMAN et al., 2018). Alguns trabalhos, principalmente fora do Brasil, vêm aplicando redes neurais artificiais como suporte à gestão de plantios da palma-de-óleo, gerando resultados satisfatórios (ALFATNI et al., 2018, 2020; CHAPMAN et al., 2018; ADIZUE et al., 2020; WATSON-HERNÁNDEZ et al., 2022).

Pensando nisso, o presente estudo teve como objetivo avaliar a eficiência das redes neurais artificiais em projetar a produção de cachos da palma-de-óleo para os cinco meses futuros, utilizando dados históricos de produção e inventário.

4.2 Material e métodos

4.2.1 Área de estudo

A área de estudo corresponde a plantios seminais comerciais da espécie *Elaeis guineensis* (Jacq.), variedade Tenera, distribuídos pelos municípios de Tailândia (sede), Tomé-Açu, Moju e Acará., localizados no estado do Pará (Figura 1). Em Tailândia, o clima é do tipo Am, tropical monçônico, com precipitação do mês mais seco inferior a 60 mm, enquanto as demais regiões Tomé-Açu, Acará e Moju apresentam clima Af, tropical sem estação seca com precipitação do mês mais seco superior a 60 mm. Ambos apresentam temperatura média do mês mais frio do ano acima de 18 °C (ALVARES et al., 2014).

O Argissolo Amarelo distrófico é o tipo de solo predominante na região, caracterizado pela profundidade variável, acidez, boa drenagem e textura variando de média a argilosa. Em outras áreas ocorre o Gleissolo háplico distrófico, este é pouco profundo, ácido, hidromórfico, de coloração acinzentada e textura arenosa (RODRIGUES et al., 2005; SANTOS et al., 2018).

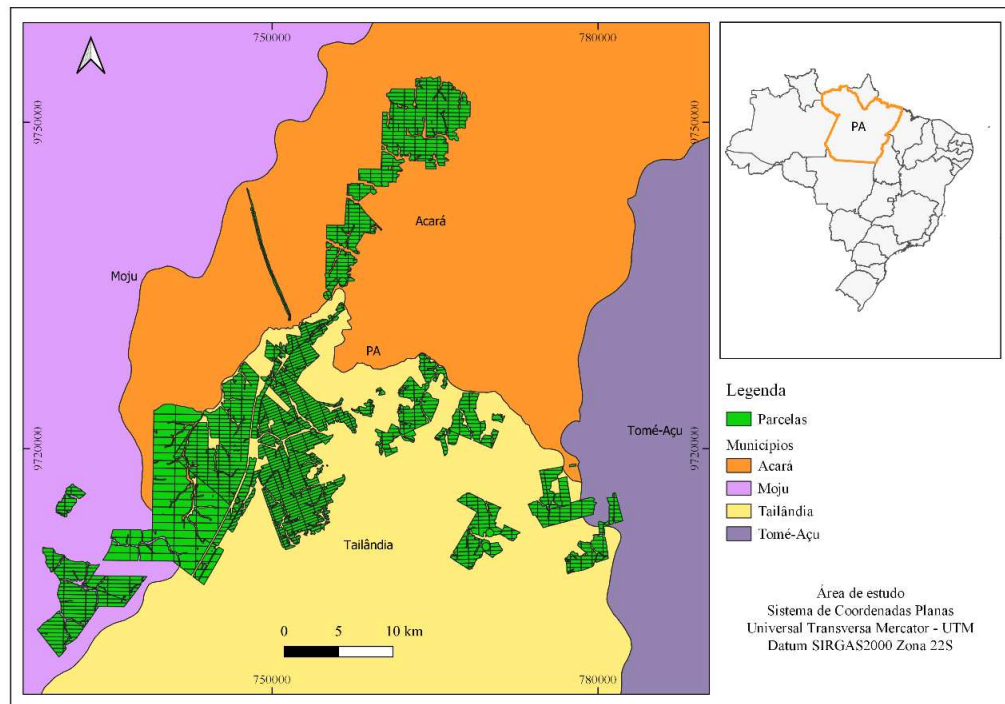


Figura 1 - Localização da área de estudo.

4.2.2 Organização da base de dados

Um intervalo mínimo de previsão igual a cinco meses foi definido para estimar a produção futura de cachos da palma-de-óleo, ou seja, estando no mês presente é realizada uma estimativa para os próximos cinco meses, utilizando como variáveis indicadoras da produção futura, dados do inventário e produção defasados em até cinco meses. Assim, a base de dados foi organizada seguindo esse princípio.

Para o primeiro mês subsequente ao último mês de coleta de dados, considerou-se os dados obtidos nos últimos cinco meses. Para o segundo mês subsequente ao último mês de coleta de dados, considerou-se os dados obtidos nos últimos quatro meses, esse mesmo raciocínio foi adotado para os próximos meses, até que fossem utilizados dados apenas da medição atual para obter as estimativas de produção referentes ao quinto mês subsequente, de tal forma que:

$$P_j = f(V_{ij-k}, V_u) \mid i \in \{1, \dots, I\}, k \in \{1, \dots, 5\}, u \in \{1, \dots, U\}$$

Em que P_j é a produção para o j -ésimo mês, V_{ij-k} é o valor da variável i no k -ésimo mês anterior ao j -ésimo mês, I é o total de variáveis que serão defasadas, k é a quantidade de defasagens, V_u é o valor da u -ésima variável de entrada não defasada.

De posse da base de dados contendo informações históricas das produções e variáveis do inventário ao longo dos últimos sete anos, foram selecionadas as parcelas que possuíam pelo menos cinco medições mensais antecedentes ao mês a ser estimado, gerando uma base final de 7765 linhas de informações.

4.2.3 Projeção da produção de cachos da palma-de-óleo

A projeção consiste na estimativa da produção futura em função da produção atual (produção observada em uma idade atual) e de quaisquer outras variáveis independentes (no caso de emprego de modelos de regressão) ou de entrada (no caso de emprego de redes neurais artificiais ou aprendizado de máquina). Nesse estudo, definiu-se como variáveis de entrada para a RNA as produções em toneladas por hectare, número de inflorescências em fecunda e antese, cachos verdes e maduros e peso médio do cacho referentes aos cinco meses anteriores ao mês a ser estimada a produção de cachos ($t\ ha^{-1}$). Além destes, as variáveis categóricas ano de plantio, fazenda, solo e material genético também foram utilizadas.

O treinamento foi realizado no software Neuroforest 4.0 (BINOTI et al, 2014). As RNAs foram configuradas conforme padrão já definido pelo software, sendo oito neurônios na camada oculta, função de ativação sigmoideal, algoritmo de treinamento resilient propagation, critério de parada 3000 ciclos e separação de 70% dos dados para treinamento e 30% para validação, para cada mês a ser estimado (1º, 2º, 3º, 4º e 5º).

4.2.4 Avaliação das redes neurais artificiais

O desempenho das redes neurais artificiais foi avaliado a partir das estatísticas correlação linear entre valores preditos e observados (1), bias (2), raiz quadrada do erro quadrático médio percentual (3), erro relativo percentual médio (4), bias percentual (5) e variância (6) (ISLAM et al., 2009; CAMPOS; LEITE, 2017) e dos gráficos representando a

dispersão dos erros relativos percentuais, relação entre valores estimados e observados e histograma da frequência dos erros relativos percentuais (4).

$$r_{y\hat{y}} = \frac{\sum_{i=1}^n (Y_i - \bar{Y}) \cdot (\hat{Y}_i - \bar{\hat{Y}})}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2 \sum_{i=1}^n (\hat{Y}_i - \bar{\hat{Y}})^2}} \quad (1)$$

$$bias = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i) \quad (2)$$

$$RQEM\% = \frac{\sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}}}{\bar{Y}} \cdot 100 \quad (3)$$

$$\overline{ER}\% = \frac{1}{n} \sum_{i=1}^n \frac{(\hat{Y}_i - Y_i) \cdot 100}{Y_i} \quad (4)$$

$$bias\% = \frac{bias}{\bar{Y}} \cdot 100 \quad (5)$$

$$Var = \frac{\sum (bias - (Y_i - \hat{Y}_i))^2}{n - 1} \quad (6)$$

Em que: Y_i é o valor observado, \hat{Y}_i é o valor estimado, \bar{Y} é a média dos valores observados, $\bar{\hat{Y}}$ é a média dos valores estimados e n é o número total de observações.

4.3 Resultados e discussão

4.3.1 Projeção da produção de cachos da palma-de-óleo

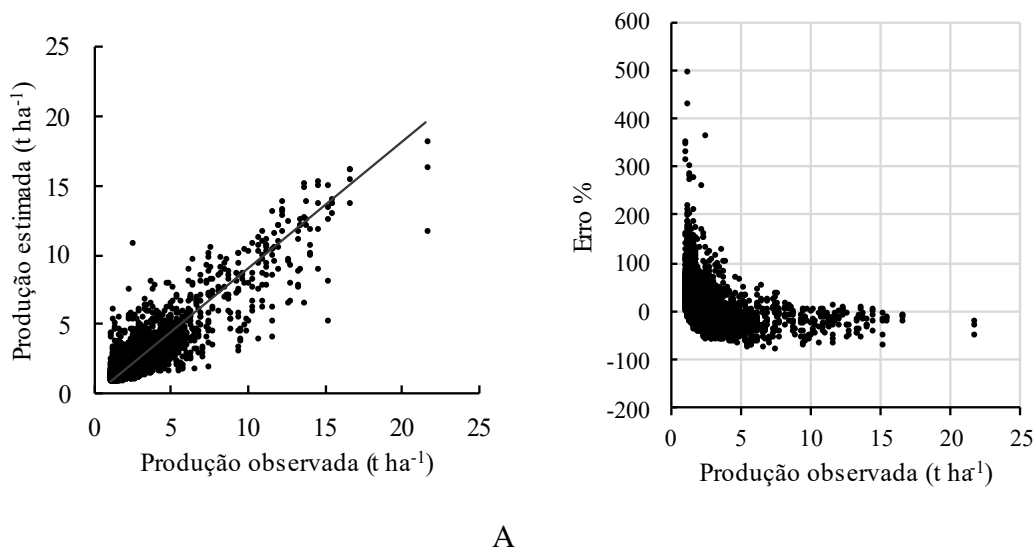
De maneira geral, a rede neuronal artificial gerou correlação entre valores observados e estimados de 0,87 e 0,78 para o treino e validação, respectivamente, bem como erros relativos médios percentuais de aproximadamente 11% para ambas as fases de processamento dos dados (Tabela 1). Confirmando a capacidade da rede neuronal artificial (RNA) em prever as produções futuras a partir de dados históricos, que exprimem o crescimento e desenvolvimento da palma-de-óleo ao longo do tempo, o que também é discutido em trabalho de Chapman et al. (2018). Os mesmos, utilizando RNAs e adotando uma janela de tempo de três anos, estimaram

o rendimento anual de cachos da palma-de-óleo a partir de dados de produção, gestão e ambientais atuais e passados (anos anteriores), obtendo uma correlação em torno de 0,6.

Tabela 1 – Estatísticas calculadas para avaliação da acurácia da RNA nas fases de treino e validação

Tipo	Bias	Bias%	RQEM%	Var	$\overline{ER\%}$	$r_{\hat{y}y}$
Treino	-0,00274	-0,10	37,47	0,96	10,54	0,87
Validação	0,02197	0,86	45,24	1,35	10,69	0,78

Os erros relativos percentuais estão distribuídos, em sua maioria, entre -100 e 100%, ocorrendo nas menores produções erros acima de 100%, principalmente para produções em torno de uma tonelada por hectare. Isso permite inferir que a RNA não consegue captar as especificidades e fatores relacionados às pequenas produções de cachos da espécie em questão (Figura 2).



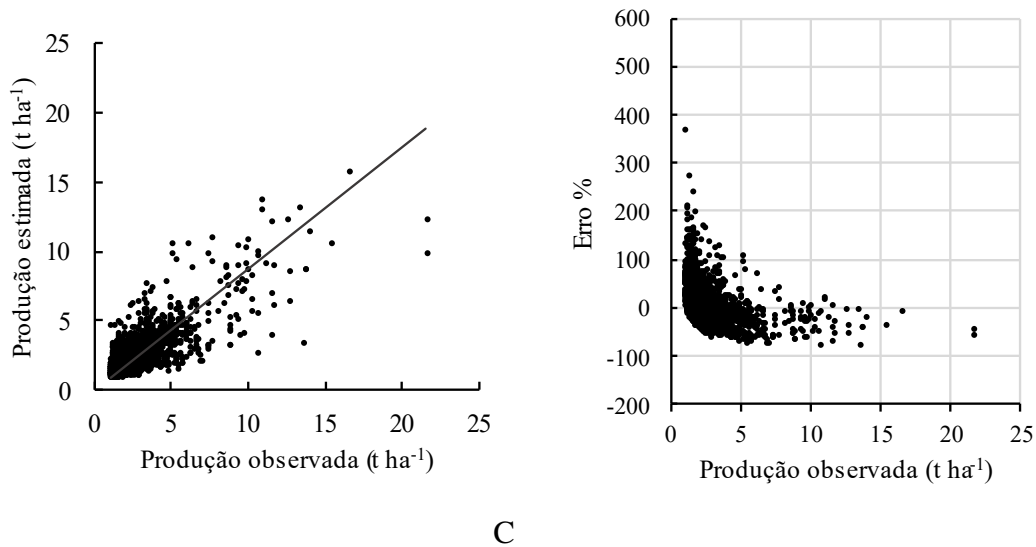


Figura 2 - Relações entre produções estimadas e observadas e entre erros relativos percentuais e produções observadas para treino (A, B) e validação (C, D) da RNA.

Avaliando os erros percentuais agrupados em classes, é possível visualizar mais de 50% dos erros concentrados nas menores classes de erro, o que é esperado em estudos de modelagem, pois independente da técnica adotada, objetiva-se diminuir o erro associado a estimativa (Figura 3) (MAESTRI et al., 2013).

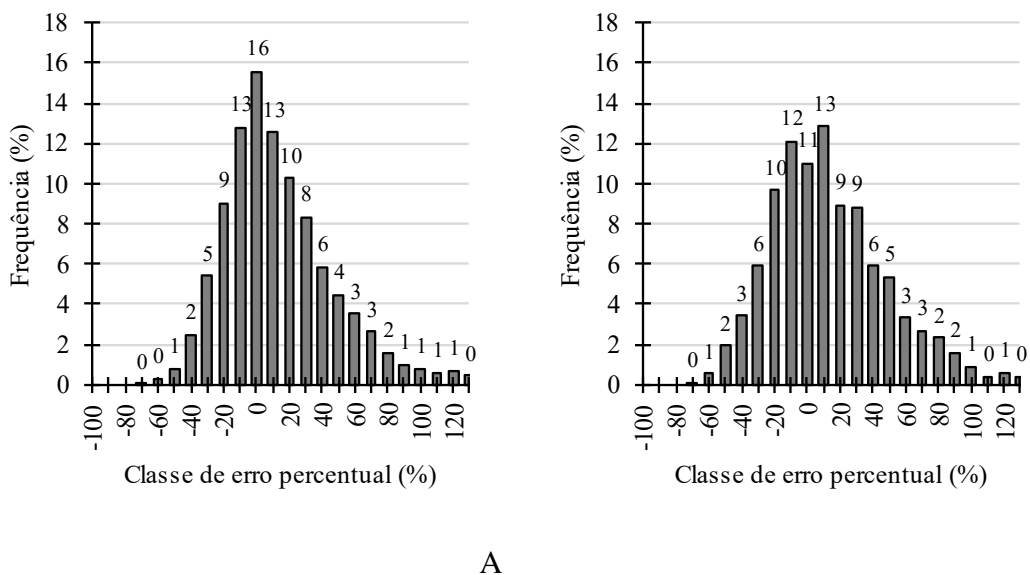
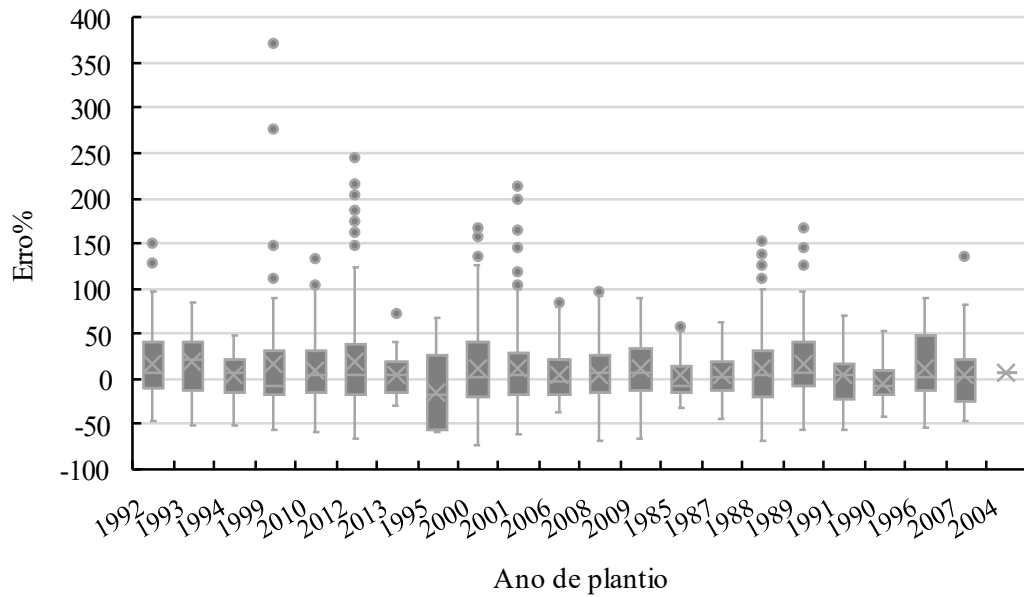


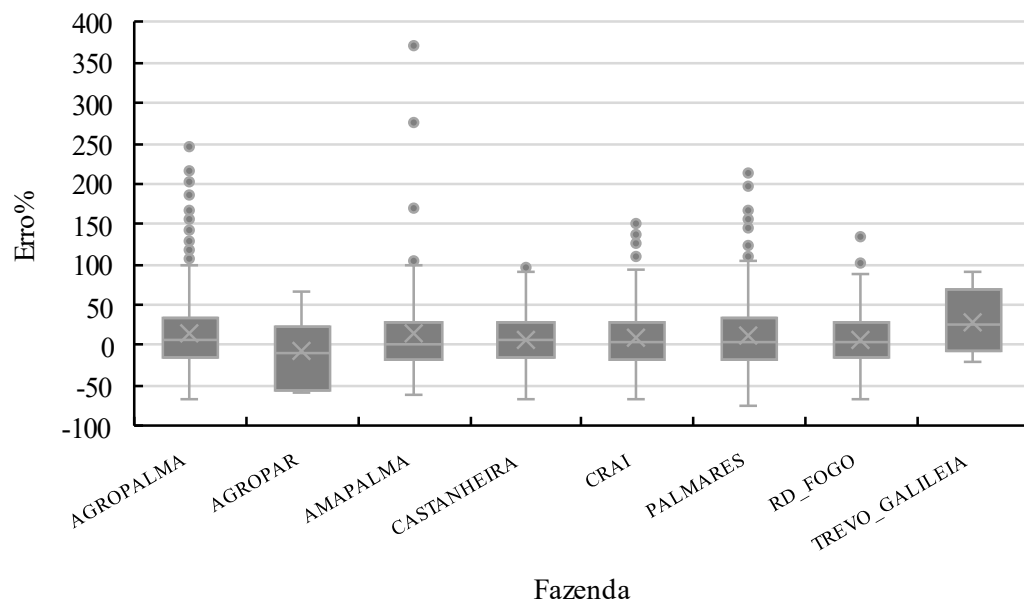
Figura 3 - Percentual de casos por classe de erro relativo percentual das estimativas de produções de cachos obtidos com os dados de treino (A) e validação (B).

Gráficos relacionando os erros percentuais com as variáveis qualitativas e quantitativas foram gerados para entender se existe alguma tendência nos dados. Nas variáveis qualitativas, a distribuição geral do erro encontra-se entre -50 e 50%, apresentando outliers em grande parte das categorias (Figura 4). As categorias de menor representatividade nos dados afastaram-se

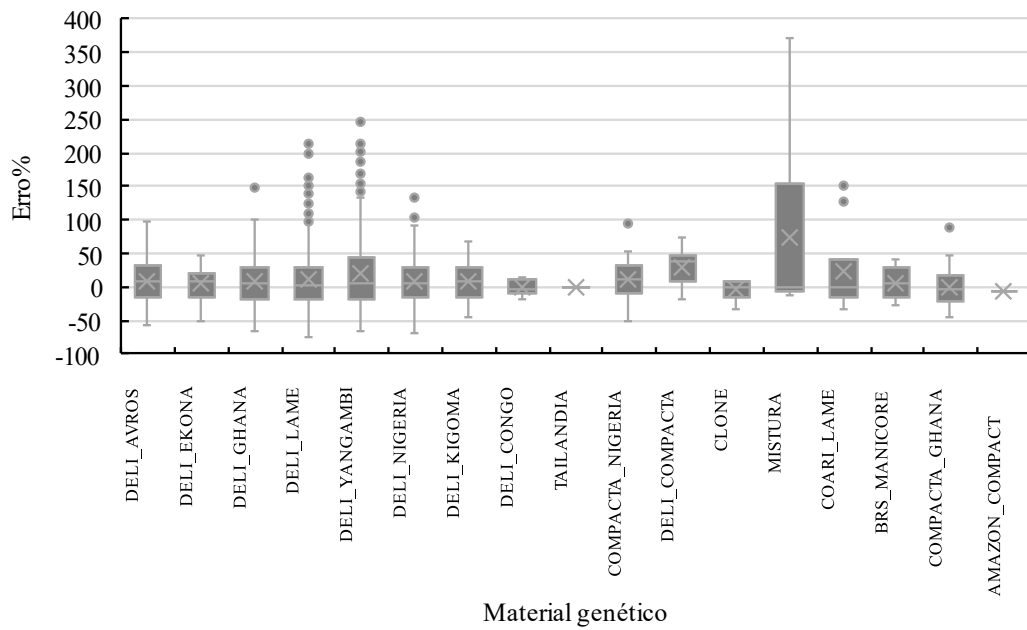
ligeiramente da tendência geral dos erros, como no ano de plantio 1995, fazendas Agropar e Trevo/Galileia e material genético Mistura. No caso do solo, isso não acontece.



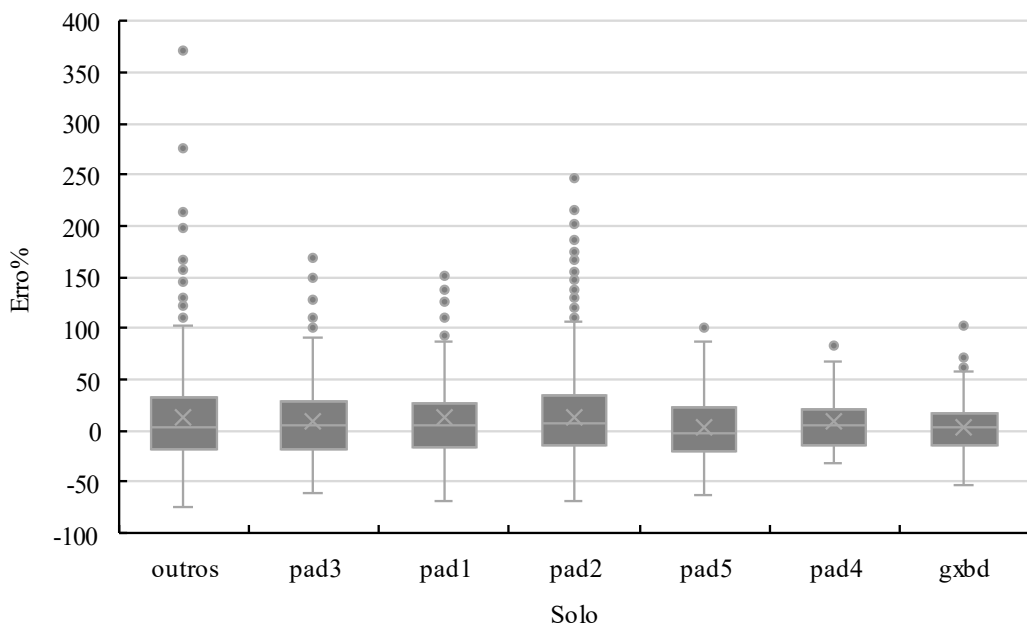
A



B



C

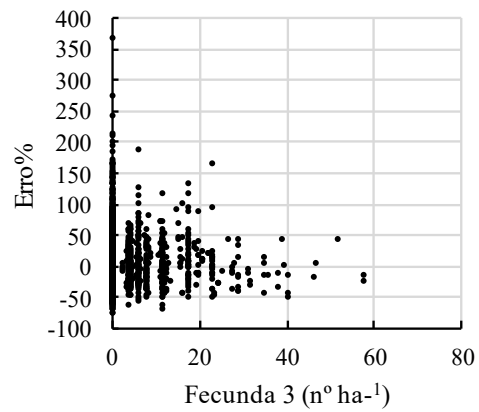
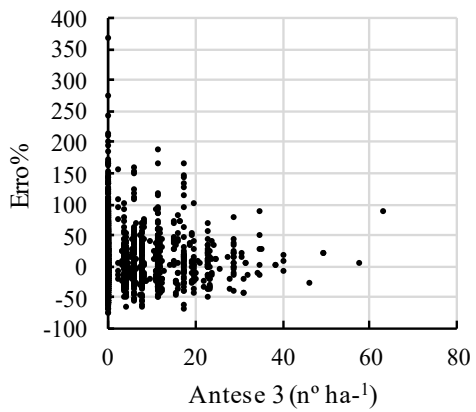
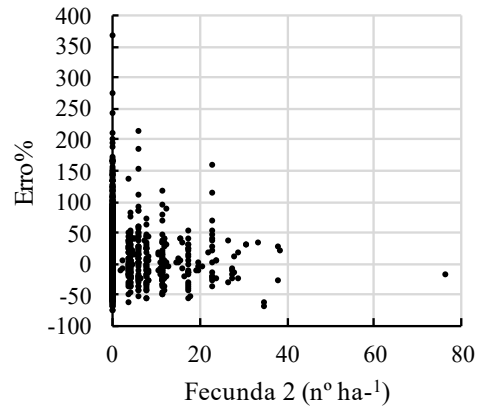
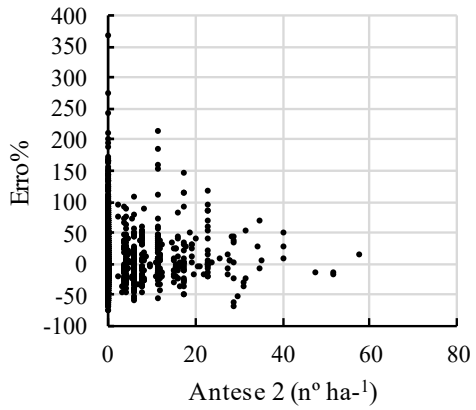
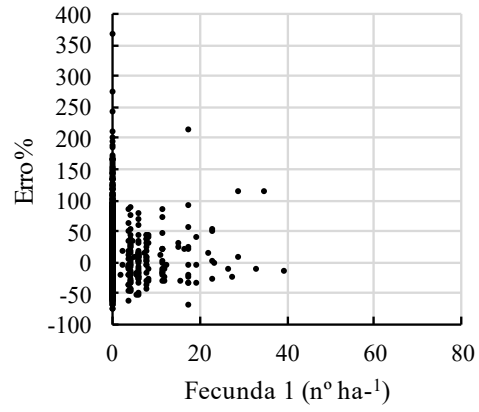
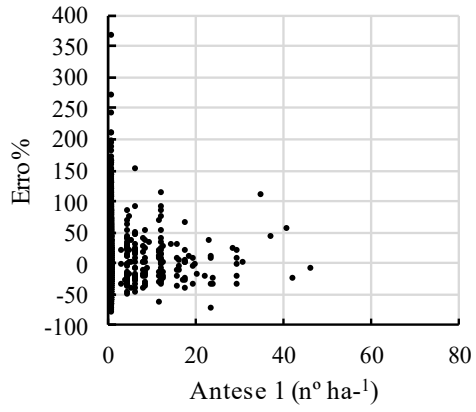


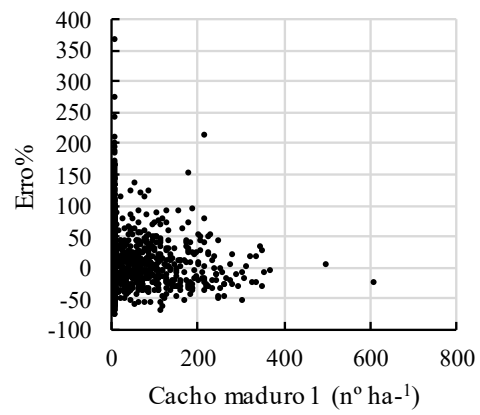
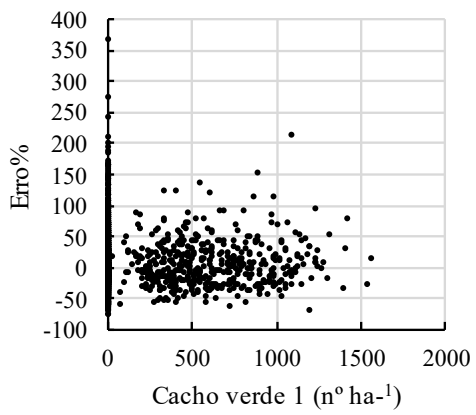
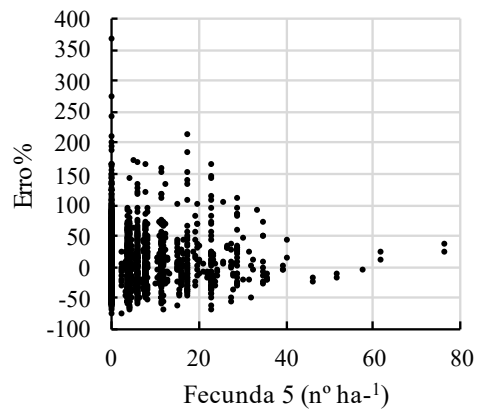
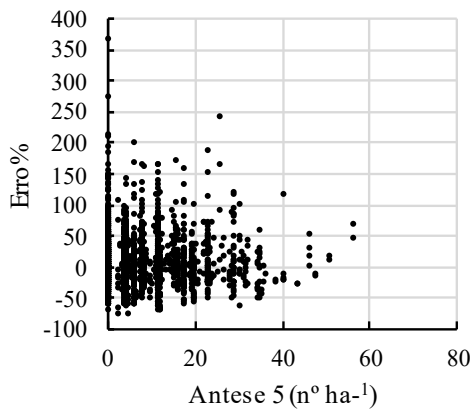
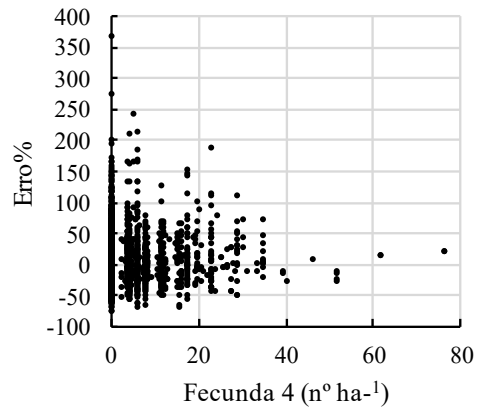
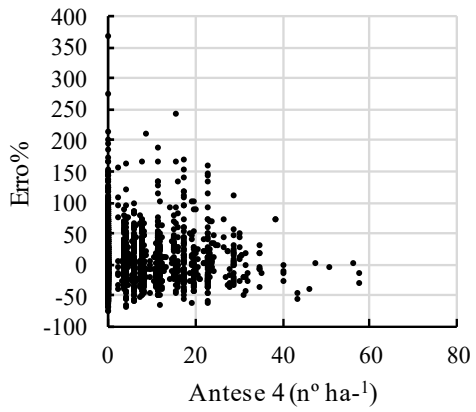
D

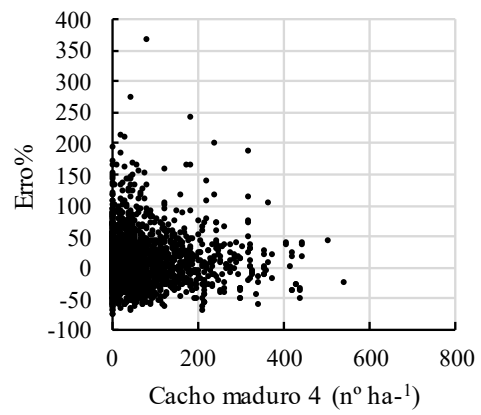
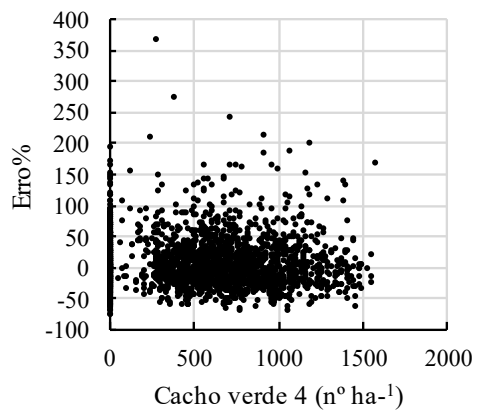
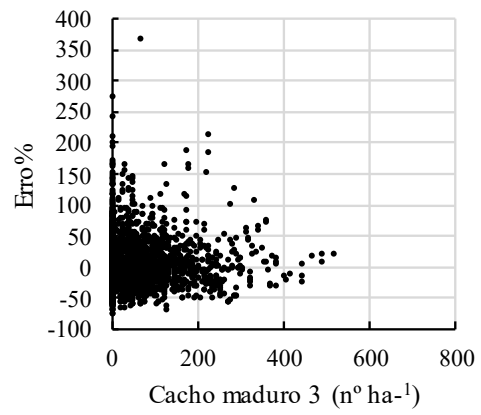
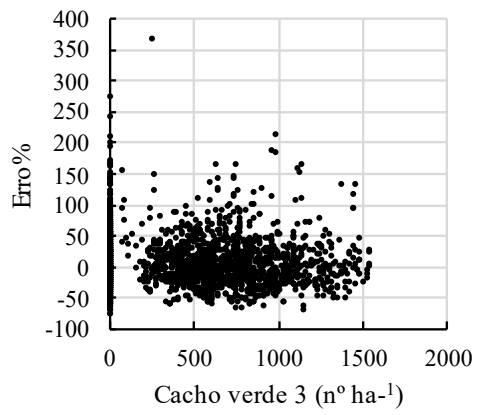
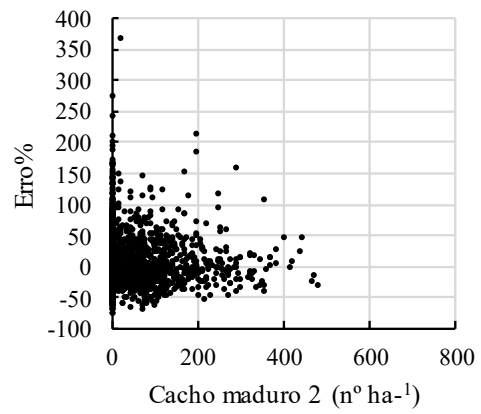
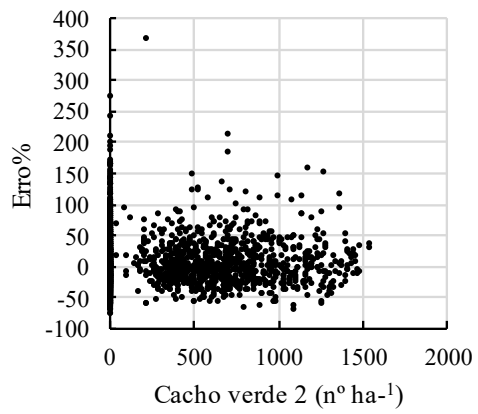
Figura 4 - Dispersão dos erros relativos percentuais avaliados por variável qualitativa na fase de validação da RNA.

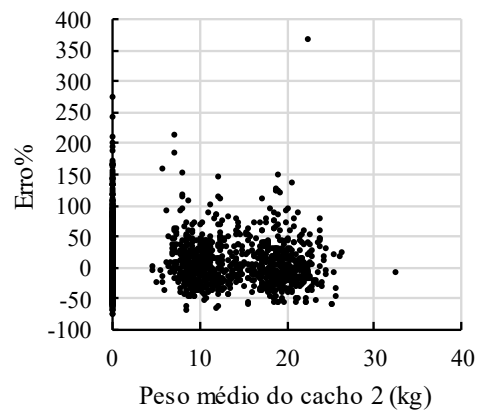
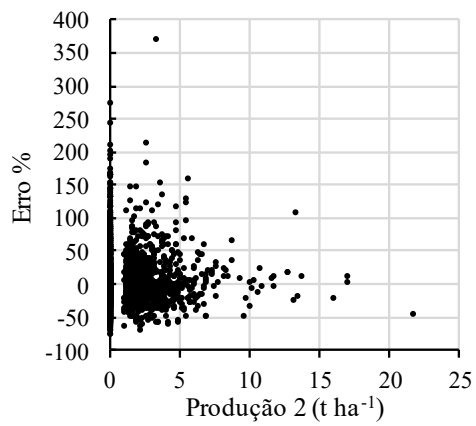
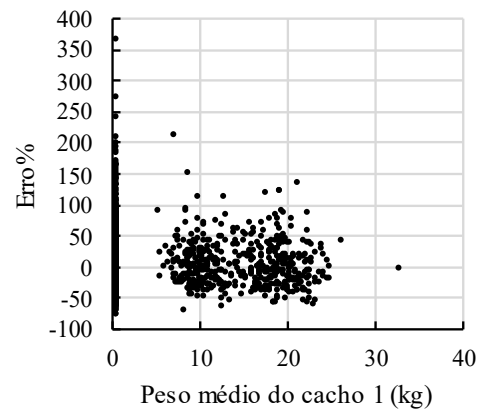
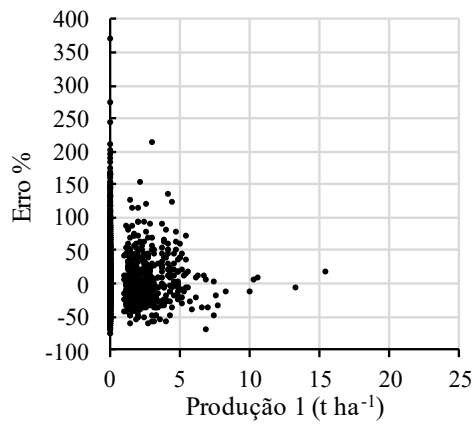
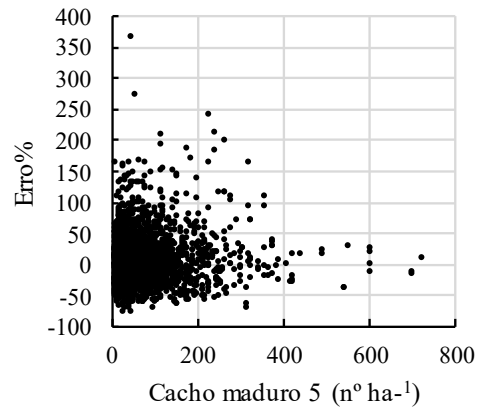
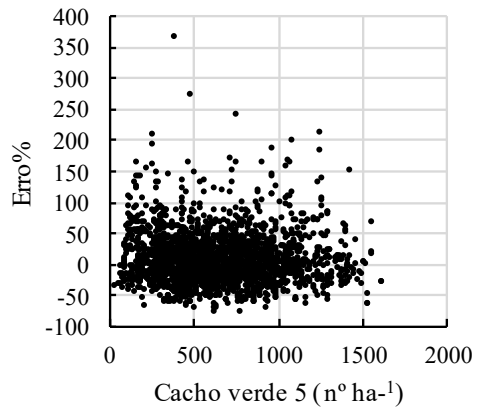
Com o aumento das variáveis quantitativas antese, fecunda, cachos maduros e produção por hectare (1, 2, 3, 4, 5), o erro tende a ser menor. Enquanto para as variáveis cachos verdes e peso médio do cacho (1, 2, 3, 4, 5) não há relação evidente com o erro (Figura 5). O processo de coleta dessas variáveis no campo, exige maior atenção de modo a evitar ao máximo erros na contabilização das inflorescências e cachos em função de sua relação com as produções futuras. Erros associados ao processo de amostragem são inevitáveis, sendo interessante desenvolver

técnicas operacionais e computacionais que reduzam o impacto desses erros na modelagem (CHAPMAN et al., 2018; MOHD NAIN et al., 2022).









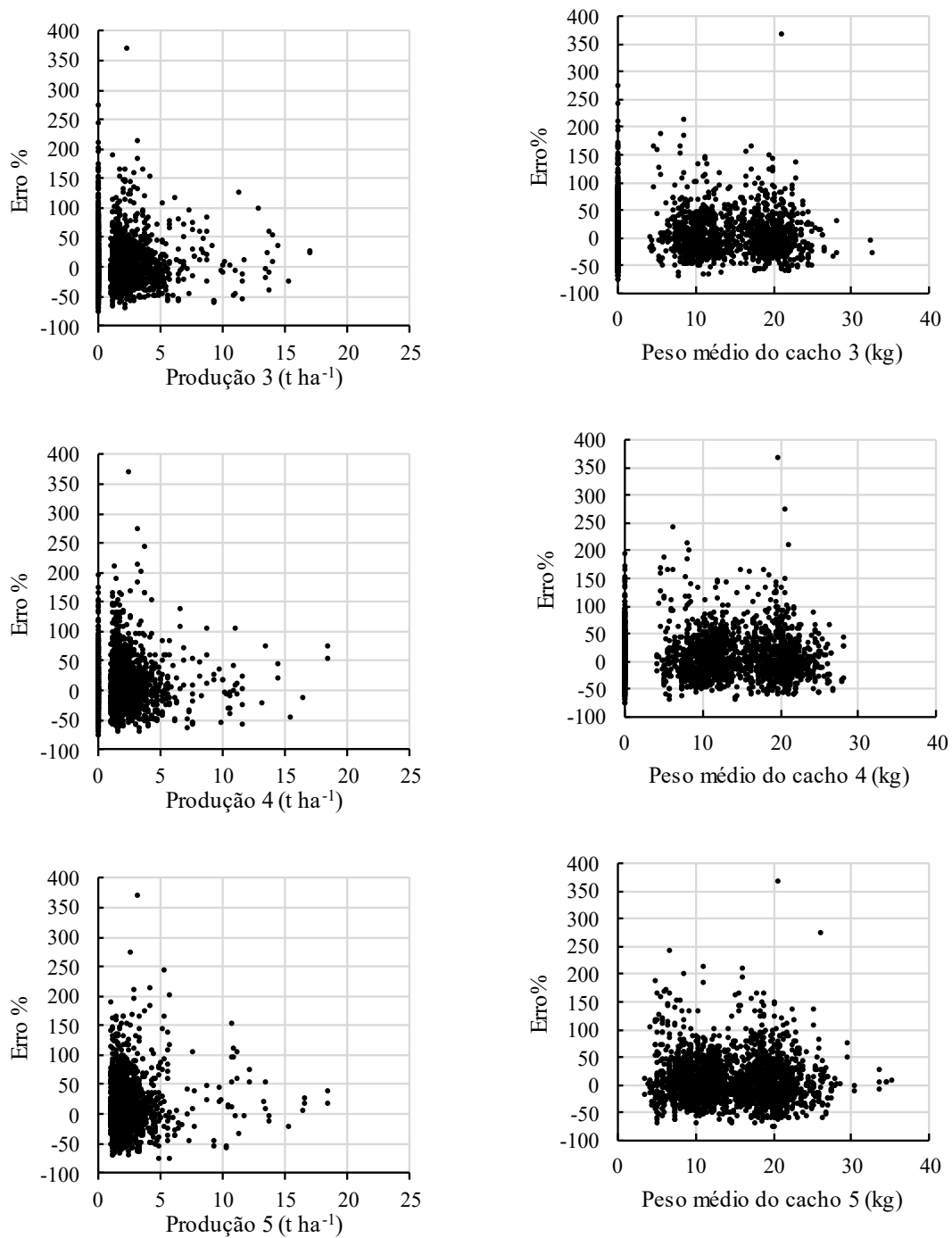


Figura 5 - Dispersão dos erros relativos percentuais avaliados por variável quantitativa na fase de validação da RNA.

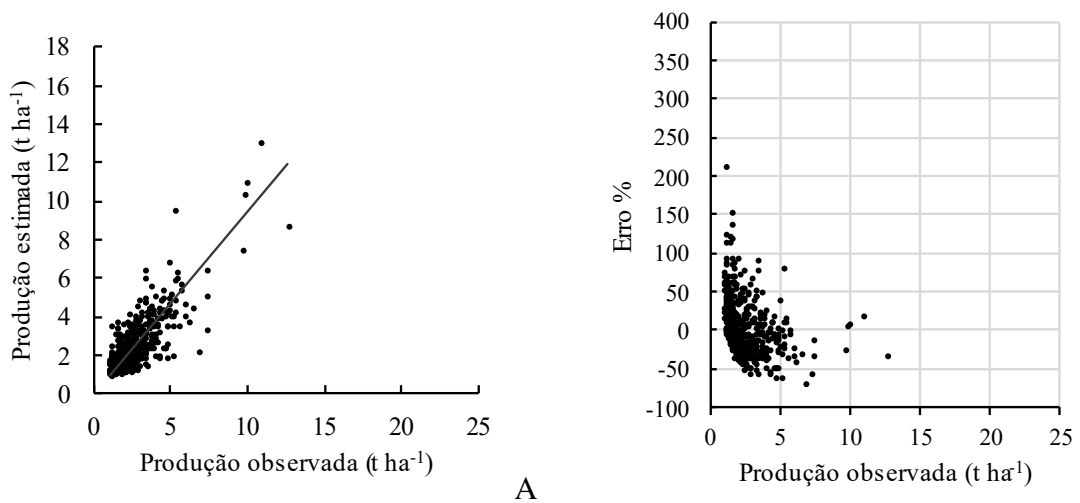
A fim de analisar o desempenho da RNA para cada mês futuro na fase de validação, foram estimados a correlação linear entre valores preditos e observados, calculadas as estatísticas de erro e gerados os gráficos correspondentes. Conforme Tabela 2, é perceptível que as estatísticas de erro ($RQEM\%$ e $\overline{ER\%}$) tendem a aumentar com o avanço dos meses. A correlação diminui, principalmente nos 4º e 5º mês, que apresentam menor correlação quando comparados aos meses anteriores. Uma possível explicação seria a disponibilidade dos dados

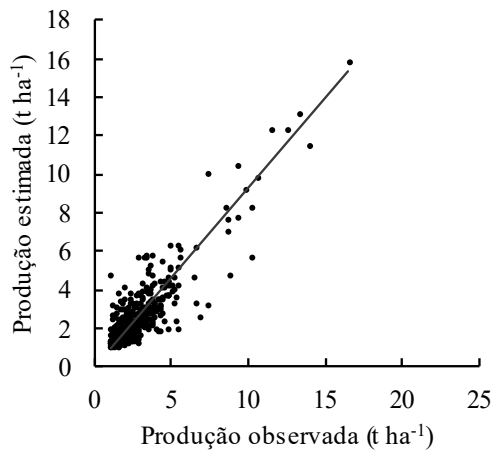
de entrada, visto que no 5º mês futuro a ser estimada a produção de cachos, estão disponíveis apenas os dados do inventário e produção referentes ao 5º mês anterior a ele e as informações dos 4º, 3º, 2º e 1º meses anteriores ainda são inexistentes.

Tabela 2 – Estatísticas calculadas para avaliação da acurácia da RNA referentes a cada produção mensal estimada na validação.

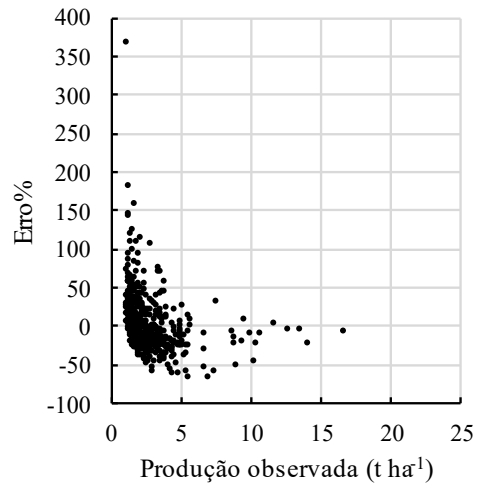
Mês	Tipo	Bias	Bias %	RQEM %	Var	$\overline{ER\%}$	$r_{y\hat{y}}$
1º	Validação	-0,007	-0,28	36,89	0,83	7,73	0,81
2º	Validação	0,005	0,21	34,22	0,82	8,61	0,88
3º	Validação	0,053	1,94	42,86	1,36	8,92	0,82
4º	Validação	0,005	0,18	49,60	1,70	13,49	0,70
5º	Validação	0,055	2,11	56,11	2,12	15,04	0,70

Em todos os meses, os erros percentuais distribuem-se, majoritariamente, entre -50 e 50%, no entanto começam a atingir valores acima de 50% nas menores produções, mesma tendência encontrada ao avaliar os dados de forma geral (Figura 6).

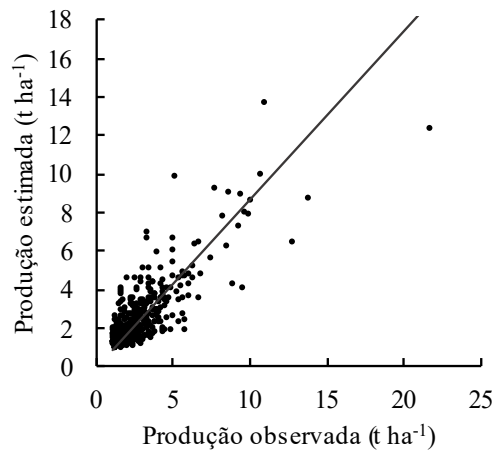




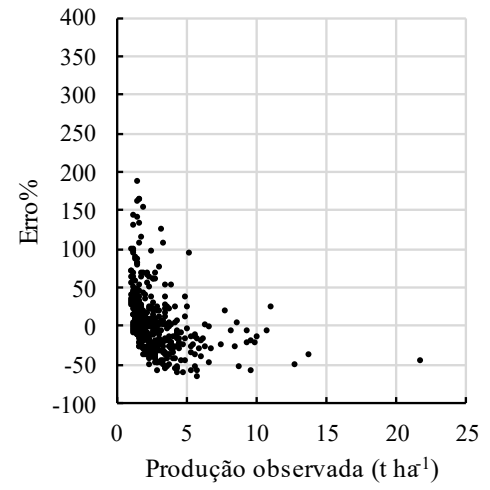
C



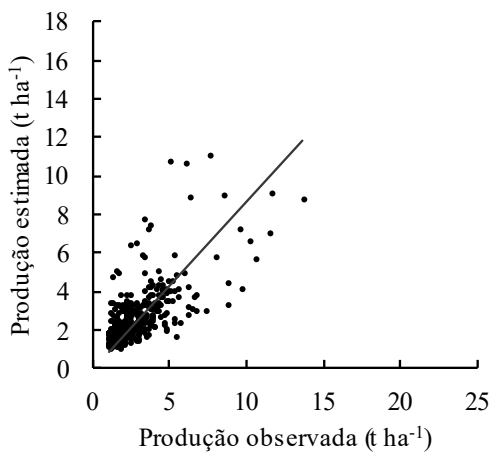
D



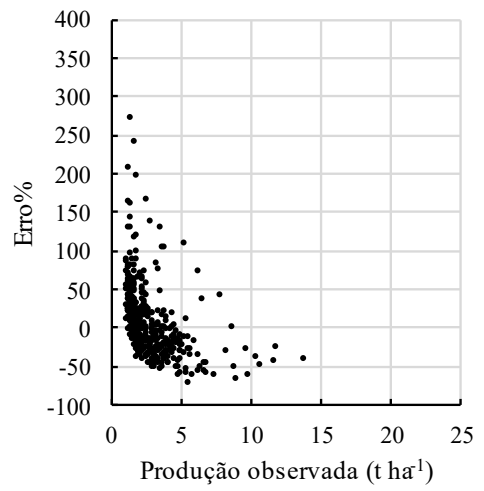
E



F



G



H

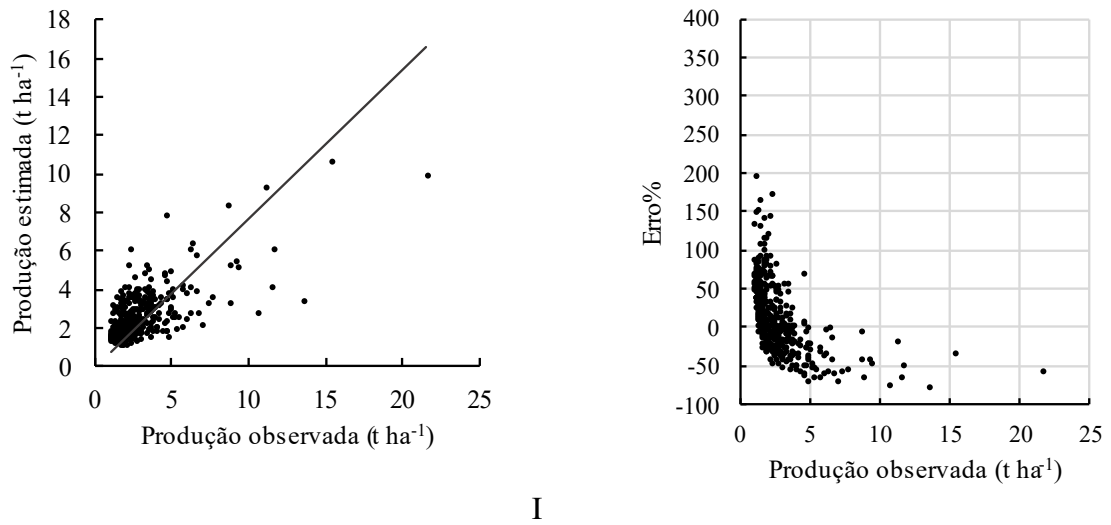


Figura 6 – Relações entre produções estimadas e observadas e entre erros relativos percentuais e produções observadas para cada mês futuro na validação. Mês 1 (A, B), Mês 2 (C, D), Mês 3 (E, F), Mês 4 (G, H), Mês 5 (I, J).

Ao gerar a produção total de cachos estimada para cada mês futuro, o erro associado ao estimado e observado é inferior a 3% (Figura 7), sendo um resultado que atende a demanda de acertar com maior precisão a produção total futura em cada mês, sob um horizonte de tempo de cinco meses.

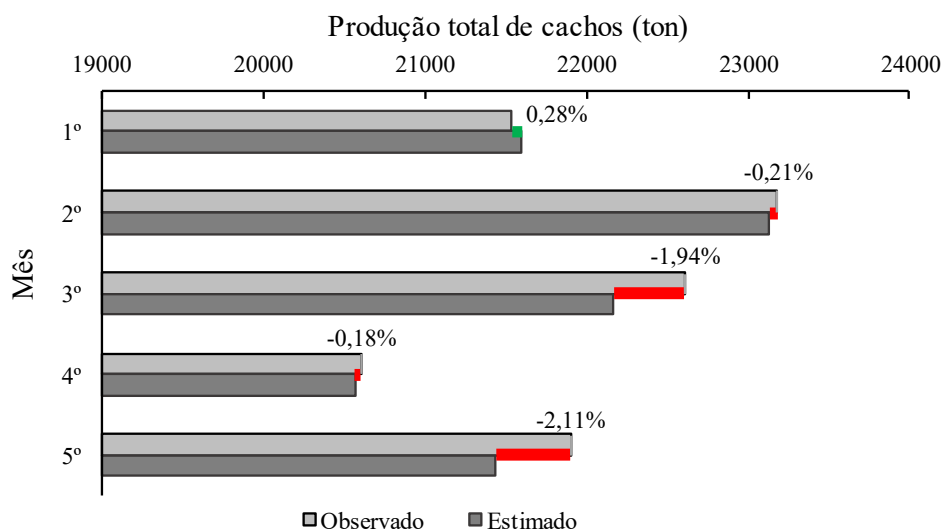


Figura 7 – Produções totais observadas e estimadas e respectivos erros percentuais para os cinco meses futuros na validação. A diferença percentual em relação à produção total observada está destacada em vermelho (subestimativa) e verde (superestimativa).

Pensando na necessidade de aplicação de uma ferramenta computacional de apoio e suporte ao gerenciamento da produção do óleo de palma, os resultados obtidos, indicam a capacidade das redes neuronais artificiais em prever o comportamento produtivo de plantios da palma-de-óleo a partir do registro dos dados de inventário e produções anteriores. A eficiência de redes neuronais artificiais em estimar produções futuras, vêm sendo comprovada, principalmente, em estudos de prognose da produção madeireira (BINOTI, 2010; ALCÂNTARA, 2015; BINOTI et al., 2015; MARTINS et al., 2015, 2016; REIS et al., 2018; SILVA, 2019). Para a palma-de-óleo é um passo importante, visto a necessidade de prever o rendimento futuro do óleo, de forma a direcionar as ações necessárias para otimização da produção.

As redes neuronais artificiais apresentam a vantagem de suportar uma diversidade de variáveis de entrada, o que é interessante para a espécie devido a influência de uma série de variáveis (ambientais, econômicas, sociais, operacionais) na produtividade (CORLEY; TINKER, 2016; BARTHEL et al., 2018). A influência de variáveis ambientais é um fator de destaque nos estudos de previsão da produção da espécie (MOHD NAIN et al., 2022), no entanto, condições sociais como direito de uso da terra para expansão da produção, o impacto do manejo e produção dos pequenos produtores, bem como questões econômicas referentes aos custos de produção e volatilidade de preços ao trabalhar com monoculturas constituem fatores a serem considerados em avaliações futuras (BARTHEL et al., 2018).

Para este estudo, a inclusão das variáveis fazenda e ano de plantio foi importante, por captar fatores específicos e ainda desconhecidos de cada região que afetam o desenvolvimento dos plantios (CHAPMAN et al., 2018). Da mesma forma, a RNA precisa diferenciar durante o treinamento os materiais genéticos e tipos de solo existentes na área, para definir algum padrão nos dados, como quais materiais genéticos e tipos de solo são mais produtivos.

As variáveis quantitativas, especificamente, informações passadas do inventário e da produção da palma-de-óleo são importantes indicadores das produções futuras em uma janela de tempo de cinco meses. A contagem de inflorescências e cachos, assim como as produções anteriores apresentam uma relação positiva com as produções a serem estimadas, portanto evitar ao máximo erros na amostragem é tarefa essencial (CHAPMAN et al., 2018; KHAN et al., 2021; MOHD NAIN et al., 2022), assim como manter uma base histórica organizada e consistida para alimentação contínua da RNA com novos dados de entrada.

As variáveis utilizadas mostraram-se capazes de explicar o comportamento produtivo futuro da palma-de-óleo. No entanto, é importante que em pesquisas futuras sejam realizados testes com novas categorias de variáveis, de modo a encontrar a combinação ótima de fatores. Mohd Nain et al. (2022), comentam que grande parte dos estudos trabalham com número limitado de recursos de entrada nos modelos, e que a adição de um número maior de variáveis poderia melhorar a precisão das estimativas, assim como ocorre em estudos envolvendo produtos agrícolas.

Estudos de previsão de rendimento do óleo de palma ainda são escassos, e não reúnem informações suficientes para determinação da melhor combinação de variáveis e quais métodos de modelagem são os mais indicados. Assim, em trabalhos futuros seria interessante testar fatores como genótipo, frutos, cachos, rendimento, reprodução, produções passadas, propriedades do solo, clima, índices de vegetação, doenças e pragas, fertilização e dados de monitoramento aéreo, utilizando diferentes métodos de inteligência artificial (KHAN et al., 2021; MOHD NAIN et al., 2022).

4.4 Conclusão

A projeção da produção de cachos da palma-de-óleo com emprego de redes neurais artificiais é eficiente e de aplicação simples. Entretanto, para que os resultados futuros tenham desempenho semelhante ou superior ao encontrado neste estudo, é necessário manter uma estrutura de base de dados histórica organizada e consistida, de forma a facilitar a atualização dos dados, à medida que forem coletados. Assim, a rede neuronal artificial será capaz de entender o padrão de comportamento das produções anteriores e aplicá-las para projetar a produção para os próximos cinco meses, gerando informações indispensáveis para o planejamento hierárquico.

4.5 Referências bibliográficas

ADIZUE, U. L.; NWANYA, S. C.; OZOR, P. A. Artificial neural network application to a process time planning problem for palm oil production. **Engineering and Applied Science Research**, v. 47, n. 2, p. 161–169, 2020.

ALCÂNTARA, A. E. M. DE. **Redes Neurais Artificiais para prognose do crescimento e da produção de povoamentos de eucalipto em Minas Gerais**, 2015. Tese (Doutorado). Viçosa: Universidade Federal de Viçosa.

ALFATNI, M. S. M.; MOHAMED SHARIFF, A. R.; BEJO, S. K.; SAAED, O. M. BEN; MUSTAPHA, A. Real-time oil palm FFB ripeness grading system based on ANN, KNN and SVM classifiers. **IOP Conference Series: Earth and Environmental Science**, v. 169, n. 1, 2018.

ALFATNI, M. S. M.; MOHAMED SHARIFF, A. R.; SAAED, O. M. BEN; ALBHBAH, A. M.; MUSTAPHA, A. Colour Feature Extraction Techniques for Real Time System of Oil Palm Fresh Fruit Bunch Maturity Grading. **IOP Conference Series: Earth and Environmental Science**, v. 540, n. 1, 2020.

ALVARES, C. A.; STAPE, L.; SENTELHAS, P. C.; GONÇALVES, L. D. M.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 2014.

AVILA-DIAZ, A.; BENEZOLI, V.; JUSTINO, F.; TORRES, R.; WILSON, A. Assessing current and future trends of climate extremes across Brazil based on reanalyses and earth system model projections. **Climate Dynamics**, v. 55, n. 5–6, p. 1403–1426, 2020. Springer Berlin Heidelberg. Disponível em: <<https://doi.org/10.1007/s00382-020-05333-z>>.

BARTHEL, M.; JENNINGS, S.; SCHREIBER, W.; SHEANE, R.; ROYSTON, S. **Study on the environmental impact of palm oil consumption and on existing sustainability standards**. 2018.

BAYAT, M.; BETTINGER, P.; HEIDARI, S.; et al. Estimation of tree heights in an uneven-aged, mixed forest in northern iran using artificial intelligence and empirical models. **Forests**, v. 11, n. 3, 2020.

BINOTI, D.; SILVA; LEITE, H. G.; LOPES, P. F.; LOPES, S. S. P., NeuroForest, Versão 4.0. Viçosa: DAP Florestal, 2014. (Software). www.neuroforest.ucoz.com

BINOTI, M. L. M. DA S. **Redes neurais artificiais para prognose da produção de povoamentos não desbastados de eucalipto**, 2010. Dissertação (Mestrado), Viçosa: Universidade Federal de Viçosa.

BINOTI, M. L. M. DA S.; LEITE, H. G.; BINOTI, D. H. B.; GLERIANI, J. M. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **Cerne**, v. 21, n. 1, p. 97–105, 2015.

CAMPOS, J.C.C.; LEITE, H.G. **Mensuração Florestal: Perguntas e Respostas**. 5. Ed.; Viçosa: Editora UFV, 2017. 636 p.

CHAPMAN, R.; COOK, S.; DONOUGH, C.; et al. Using Bayesian networks to predict future yield functions with data from commercial oil palm plantations: A proof of concept analysis. **Computers and Electronics in Agriculture**, v. 151, n. October 2017, p. 338–348, 2018.

CORLEY, R. H. V; TINKER, P. B. **The Oil Palm**. BlackWell Science Ltd, 2016.

DEY, S.; REANG, N. M.; DAS, P. K.; DEB, M. A comprehensive study on prospects of economy, environment, and efficiency of palm oil biodiesel as a renewable fuel. **Journal of Cleaner Production**, v. 286, p. 124981, 2021. Elsevier Ltd. Disponível em: <<https://doi.org/10.1016/j.jclepro.2020.124981>>.

ISLAM, M. N.; KURTTILA, M.; MEHTÄTALO, L.; HAARA, A. Analyzing the effects of inventory errors on holding-level forest plans: The Case of measurement error in the basal area of the dominated tree species. **Silva Fennica**, v. 43, n. 1, p. 71–85, 2009.

KHAN, N.; KAMARUDDIN, M. A.; SHEIKH, U. U.; YUSUP, Y.; BAKHT, M. P. Oil palm and machine learning: Reviewing one decade of ideas, innovations, applications, and gaps. **Agriculture (Switzerland)**, v. 11, n. 9, p. 1–26, 2021.

LEITE, H. G.; ALCÂNTARA, A. E. M.; BINOTI, D. H. B.; OLIVEIRA NETO, R. R.; SILVA, M. L. M. Comparação entre predição e projeção da distribuição de diâmetros de povoamentos de Eucalipto submetidos a desbastes. **Revista Árvore**, v. 37, n. 2, p. 321–328, 2013.

MAESTRI, R.; SANQUETTA, C. R.; SCOLFORO, J. R.; MACHADO, S. DO A.; CORTE, A. P. D. Modelagem Do Crescimento Florestal Considerando Variáveis Do Ambiente: Revisão. **Scientia Agraria**, v. 14, n. 3, 2013.

MARTINS, E. R.; BINOTI, M. L. M. D. S.; LEITE, H. G.; BINOTI, D. H. B.; DUTRA, G. C. Configuração de redes neurais artificiais para prognose da produção de povoamentos clonais de eucalipto. **Revista Brasileira de Ciências Agrárias**, v. 10, n. 4, p. 532–537, 2015.

MARTINS, E. R.; BINOTI, M. L. M. DA S.; LEITE, H. G.; BINOTI, D. H. B.; DUTRA, G. C. Configuração de redes neurais artificiais para estimação do afilamento do fuste de árvores de eucalipto. **Revista Brasileira de Ciências Agrárias**, v. 11, n. 1, p. 33–38, 2016.

MBA, O. I.; DUMONT, M. J.; NGADI, M. Palm oil: Processing, characterization and utilization in the food industry - A review. **Food Bioscience**, v. 10, p. 26–41, 2015.

MOHD NAIN, F. N.; AHAMED HASSAIN MALIM, N. H.; ABDULLAH, R.; et al. A Review of an Artificial Intelligence Framework for Identifying the Most Effective Palm Oil Prediction. **Algorithms**, v. 15, n. 6, p. 218, 2022.

PIRKER, J.; MOSNIER, A.; KRAXNER, F.; HAVLÍK, P.; OBERSTEINER, M. **What are the limits to oil palm expansion?** **Global Environmental Change**, v. 40, p. 73–81, 2016. Elsevier Ltd.

REIS, L. P.; SOUZA, A. L. DE; REIS, P. C. M. DOS; et al. Diameter structure in a community of shrub-tree Caatinga, municipality of Floresta, state of Pernambuco, Brazil. **Floresta**, v. 48, n. 1, p. 133–142, 2018.

RODRIGUES, T. E.; SILVA, J. M. L.; SILVA, B. N. R.; VALENTE, M. A.; GAMA, J. R. N. F.; SANTOS, E. S.; ROLLIM, P. A. M.; PONTE, F. C. **Caracterização e Classificação dos Solos do Município de Tailândia, Estado do Pará**. Belém: Embrapa Amazônia Oriental, 2005.

SANTOS, H. G. DOS; JACOMINE, P. K. T.; ANJOS, L. H. C. DOS; OLIVEIRA, V. A.; LUMBRERAS, J. F.; COELHO, M. R.; ALMEIDA, J. A.; ARAÚJO FILHO, J. C.; OLIVEIRA, J. B.; CUNHA, T. J. F. **Sistema brasileiro de classificação de solos**. Brasília: Embrapa solos, 2018.

SEMBIRING, P.; MAWENGKANG, H.; SADYADHARMA, H. Mixed Integer Linear Programming model for Crude Palm Oil Supply Chain Planning. **IOP Conference Series: Materials Science and Engineering**. Anais..., v. 300, p.2–7, 2017. Medan.

SILVA, P. E. DA; SANTOS E SILVA, C. M.; SPYRIDES, M. H. C.; ANDRADE, L. DE M. B. Precipitation and air temperature extremes in the Amazon and northeast Brazil. **International Journal of Climatology**, v. 39, n. 2, p. 579–595, 2019.

SILVA, W. D. M. **Redes neurais artificiais como ferramenta para prognose de crescimento e melhoramento genético florestal**, 2019. Tese (Doutorado), Jaboticabal: Universidade Estadual Paulista.

WATSON-HERNÁNDEZ, F.; GÓMEZ-CALDERÓN, N.; SILVA, R. P. DA. Oil Palm Yield Estimation Based on Vegetation and Humidity Indices Generated from Satellite Images and Machine Learning Techniques. **AgriEngineering**, v. 4, n. 1, p. 279–291, 2022.

5 CONCLUSÃO GERAL

Redes neuronais artificiais de configuração simples são eficientes para modelar a produção (atual e futura) de cachos da palma-de-óleo. Para a predição da produção, variáveis relacionadas aos estratos de gerenciamento dos plantios, tipos de solo, materiais genéticos e o número de cachos maduros por hectare são importantes, assim como elementos climáticos, especialmente em situações de elevadas variações ao longo do ano.

O emprego da rede neuronal permite antecipar a expectativa de produção, por parcela e total, em até cinco meses antes da colheita, utilizando dados históricos de produção e inventário. No caso da produção total o erro de projeção é insignificante.

Os resultados aqui obtidos comprovam a possibilidade de adoção de técnicas computacionais como auxílio ao processo de tomada de decisão na produção de óleo de palma.