

# RESSALVA

Atendendo solicitação do(a)  
autor(a), o texto completo desta tese  
será disponibilizado somente a partir  
de 09/05/2025.

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP  
CÂMPUS DE JABOTICABAL**

**MODELOS DE MACHINE LEARNING PARA ESTIMAÇÃO DE  
PRODUTIVIDADE DE SOJA E EUCALIPTO NO CERRADO  
BRASILEIRO**

**Valter Barbosa dos Santos**

Engenheiro Agrônomo

**2024**

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP  
CÂMPUS DE JABOTICABAL**

**MODELOS DE MACHINE LEARNING PARA ESTIMAÇÃO DE  
PRODUTIVIDADE DE SOJA E EUCALIPTO NO CERRADO  
BRASILEIRO**

**Valter Barbosa dos Santos**

**Orientador: Prof. Dr. Glauco de Souza Rolim**

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Doutor em Agronomia (Ciência do Solo).

S237m

Santos, Valter Barbosa dos

Modelos de machine learning para estimação de produtividade de soja e eucalipto no Cerrado brasileiro / Valter Barbosa dos Santos. -- Jaboticabal, 2024

78 p.

Tese (doutorado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientador: Glauco de Souza Rolim

1. Agricultura. 2. Floresta. 3. Agrometeorologia. 4. Machine learning. I. Título.

**CERTIFICADO DE APROVAÇÃO**

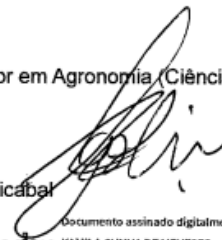
TÍTULO DA TESE: MODELOS DE MACHINE LEARNING PARA ESTIMAÇÃO DE PRODUTIVIDADE DE SOJA E EUCALIPTO NO CERRADO BRASILEIRO

**AUTOR: VALTER BARBOSA DOS SANTOS**

**ORIENTADOR: GLAUCO DE SOUZA ROLIM**

Aprovado como parte das exigências para obtenção do Título de Doutor em Agronomia (Ciência do Solo), pela Comissão Examinadora:

Prof. Dr. GLAUCO DE SOUZA ROLIM (Participação Virtual)  
Departamento de Engenharia e Ciências Exatas / FCAV UNESP Jaboticabal



Profa. Dra. KAMILA CUNHA DE MENESES (Participação Virtual)  
Coordenação de Zootecnia / Universidade Federal do Maranhão (L

gov.br KAMILA CUNHA DE MENESES  
Data: 11/09/2024 12:25:10-0300  
Verifique em <https://validar.itl.gov.br>

Documento assinado digitalmente

Prof. Dr. ALEXANDRE DAL PAI (Participação Virtual)  
Departamento de Bioprocessos e Biotecnologia / FCA UNESP Botucatu

gov.br ALEXANDRE DAL PAI  
Data: 16/09/2024 18:27:28-0300  
Verifique em <https://validar.itl.gov.br>

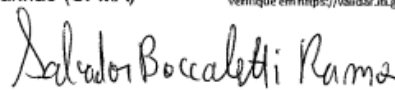
Documento assinado digitalmente

Prof. Dr. GUSTAVO ANDRÉ DE ARAÚJO SANTOS (Participação Virtual)  
Centro de Ciências de Chapadinha / Universidade Federal do Maranhão (UFMA)

gov.br GUSTAVO ANDRÉ DE ARAÚJO SANTOS  
Data: 11/09/2024 14:38:57-0300  
Verifique em <https://validar.itl.gov.br>

Documento assinado digitalmente

Prof. Dr. SALVADOR BOCCALETTI RAMOS (Participação Virtual)  
Departamento de Ciências Exatas / FCAV UNESP Jaboticabal



Jaboticabal, 09 de maio de 2024

## DADOS CURRICULARES DO AUTOR

**Valter Barbosa dos Santos**- Filho de José Carlos dos Santos e Severina Barbosa (*In memoriam*). Nasceu em São José de Ribamar, Maranhão, no dia 22 de fevereiro de 1990. Técnico em Agropecuária pela Escola Agrotécnica Federal do Maranhão (2009), cursou engenharia Agrônômica na Universidade Estadual do Maranhão-UEMA, câmpus Paulo VI, de São Luís-MA, de 2010 a 2017. Bolsista de iniciação científica pela PIBIC/UEMA em 2014-2015, atuando principalmente nos seguintes temas: manejo de cultivos agrícolas, potássio, adubação orgânica, agricultura familiar, Fertilidade do Solo. Trabalhou como servidor público, concursado no cargo de Fiscal Ambiental em 2013-2018. Em março de 2018, ingressou no Curso de Mestrado em Agronomia (Ciência do Solo), na Faculdade de Ciências Agrárias e Veterinárias – UNESP, desenvolvendo pesquisa sobre modelagem e inteligência artificial para aplicação no setor agrícola. Em março de 2020, foi aprovado no Curso de Doutorado em Agronomia (Ciência do Solo), mantendo a linha pesquisa em modelagem e Inteligência artificial. É integrante do grupo de pesquisa: “Group of Agrometeorological Studies” (GAS), da Unesp – Câmpus de Jaboticabal.

***Debaixo do céu há momento para tudo, e tempo certo para cada coisa.***

***Eclesiastes 3,1***

## **DEDICO**

A Deus, pela dádiva da vida pois sem ele eu não teria traçado o meu caminho e feito a minha escolha pela Agronomia.

Aos meus pais, José Carlos e Severina Barbosa  
(*In memoriam*) pelo apoio e confiança, mesmo distantes. Ao meu filho pela companhia nessa jornada.

## **OFEREÇO**

À minha família, pelo incentivo e força nessa caminhada.

E ao professor Glauco Rolim pelo excepcional papel de orientador que desempenha na Unesp-Jaboticabal.

## AGRADECIMENTOS

Agradeço primeiramente a **Deus**, pela dádiva da vida pois sem ele eu não teria traçado o meu caminho e feito a minha escolha pela Agronomia.

A minha **família** pelo amor incondicional, especialmente meus pais e meus irmãos, acreditando, apoiando e confiando em mim, aos meus sobrinhos e a toda família Moreno e Barbosa dos Santos.

Ao meu filho Samuel pelo amor e por me acompanhar nessa jornada.

Ao meu orientador professor Glauco de Souza Rolim, por sua amizade, dedicação pelos seus ensinamentos que me possibilitou realizar mais essa etapa da minha vida.

A professora Dr<sup>a</sup>. Ana Maria Silva de Araújo, pelos ensinamentos, amizade, disposição, e paciência, por ter acreditado no meu potencial desde 2014 quando começamos a trabalhar juntos, me despertando o lado científico das ciências agrônomicas através das orientações em projetos de iniciação científica entre os anos de 2014 a 2015.

Ao professor Gener Tadeu Pereira e ao meu amigo José Reinaldo Moraes membros da banca de qualificação pelas importantes sugestões.

Ao Grupo de Pesquisa em Agrometeorologia da Unesp – GAS, pelo recebimento no grupo e pelos conhecimentos compartilhados.

Aos funcionários do Departamento de Ciências Exatas, Maria José Servidone Trizólio, Shirley Aparecida Martineli de Sousa, Adriana Elisabete Takakura, por me receberem bem no departamento, e pelo carinho.

Ao programa de Pós-Graduação em Agronomia (Ciência do Solo), pela oportunidade em cursar o mestrado.

A todos os meus amigos que contribuíram direta e indiretamente nessa caminhada.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

## SUMÁRIO

<b>RESUMO.....</b>	<b>x</b>
<b>ABSTRACT.....</b>	<b>xi</b>
<b>LISTA DE ABREVIATURAS E SIGLAS.....</b>	<b>xii</b>
<b>LISTA DE FIGURAS.....</b>	<b>xiii</b>
<b>LISTA DE TABELAS.....</b>	<b>xiv</b>
<b>CAPÍTULO 1 – Considerações gerais.....</b>	<b>16</b>
1.1 Introdução .....	16
1.2 Revisão de Literatura .....	18
1.2.1 Panorama da Soja no Cerrado .....	18
1.2.2 Panorama do Eucalipto no Cerrado.....	19
1.2.3 Modelagem da produtividade.....	21
1.3 Objetivo geral .....	23
<b>REFERÊNCIAS.....</b>	<b>24</b>
<b>CAPÍTULO 2: Algoritmos de aprendizado de máquina para previsão de produtividade de soja no Cerrado brasileiro. ....</b>	<b>31</b>
<b>RESUMO.....</b>	<b>31</b>
2.1 Introdução .....	32
2.2 Material e métodos.....	33
2.2.1 Localização e caracterização da área de estudo .....	33
2.2.2 Conjunto de dados .....	36
2.2.3 Modelos de machine learning .....	38
2.2.3.1 Random Forest – RF .....	38
2.2.3.2 RNA – Multilayer Perceptrom – RNA – MLP .....	38
2.2.3.3 Suport vector machines – SVM.....	39
2.2.4 Avaliação do modelo .....	39
2.3 Resultados e Discussão .....	40
2.4 Conclusão .....	47
<b>REFERÊNCIAS.....</b>	<b>47</b>
<b>Capítulo 3 – Modelos de machine learning para estimações do volume de eucalipto no Cerrado Brasileiro a partir de dados climáticos sazonais....</b>	<b>50</b>
<b>RESUMO.....</b>	<b>50</b>
3.1 introdução.....	51
3.2 Material e métodos.....	53
3. 2.1 Caracterização da área de estudo .....	53
3.2.2. Dados meteorológicos.....	54
3.2.3 Inventario florestal .....	59
3.2.4 Regressão Linear Múltipla com seleção de variáveis por stepwise backwards .....	63
3.2.5 Otimização de parâmetros dos modelos de machine learning.....	67
3.2.6 Métricas de avaliação dos modelos .....	68
3.3 Resultados e Discussão .....	68
3.4. Conclusões.....	74
<b>Referências.....</b>	<b>75</b>

## MODELOS DE MACHINE LEARNING PARA ESTIMAÇÃO DE PRODUTIVIDADE DE SOJA E EUCALIPTO NO CERRADO BRASILEIRO

**RESUMO** – As mudanças ocorridas no Cerrado promoveram grandes desafios para o bioma. Em geral, a conservação do Cerrado tem sido conduzida por meio de políticas nacionais e locais. Essas políticas consideram a variação cultural e socioeconômica entre os municípios do Cerrado proporcionando avanços tecnológicos tanto na área florestal quanto na agricultura, tornando o Brasil líder mundial na produção de soja. Buscando reduzir a pressão do desmatamento e a manutenção da biodiversidade, ocorreu a introdução do eucalipto no Cerrado. Dessa forma, avaliou-se diferentes modelos de machine learning para estimação de produtividade de soja para o sul do Maranhão, com até um mês antecedência. Os resultados mostraram que o algoritmo Random Forest - RF atinge a maior precisão e acurácia, com  $R^2$  de 0,81, RMSE de  $176,93 \text{ kg ha}^{-1}$  e tendência (EME) de  $1,99 \text{ kg ha}^{-1}$ . Por outro lado, o algoritmo Suport vector machine kernel RBF - SVM\_RBF apresentou o menor desempenho com  $R^2$  de 0,74, RMSE de  $213,58 \text{ kg ha}^{-1}$  e EME de  $15,06 \text{ kg ha}^{-1}$ . Em um segundo estudo buscamos estimar o volume de madeira de eucalipto no cerrado brasileiro utilizando técnicas de machine learning e apenas dados climáticos como entrada dos modelos, abrangendo diferentes idades de crescimento em dois períodos do ano entre janeiro e junho e entre julho e dezembro. Os modelos apresentaram ótimos resultados na estimativa do volume de madeira. O modelo Random Forest apresentou as melhores métricas durante o treinamento e teste com,  $R^2= 0,93$  e  $\text{RMSE} = 18,36 \text{ m}^3\text{ha}^{-1}$  para o modelo janeiro-junho e  $R^2= 0,92$  e  $\text{RMSE} = 19,52 \text{ m}^3\text{ha}^{-1}$  para o modelo de julho-dezembro.

**Palavras chaves:** agricultura, floresta, agrometeorologia, machine learning

## MACHINE LEARNING MODELS FOR ESTIMATING SOYBEAN AND EUCALYPTUS PRODUCTIVITY IN THE BRAZILIAN CERRADO

**ABSTRACT** – The changes that occurred in the Cerrado promoted major challenges for the biome. In general, Cerrado conservation has been driven through national and local policies. These policies consider the cultural and socioeconomic variation between the municipalities of the Cerrado, providing technological advances in both forestry and agriculture, making Brazil a world leader in soybean production. Seeking to reduce the pressure of deforestation and maintain biodiversity, eucalyptus was introduced into the Cerrado. In this way, different ML models were evaluated to predict soybean productivity for the south of Maranhão, up to one month in advance. The results showed that the RF algorithm achieves the highest precision and accuracy, with  $R^2$  of 0.81, RMSE of 176.93 kg ha<sup>-1</sup> and trend (EME) of 1.99 kg ha<sup>-1</sup>. On the other hand, the SVM\_RBF algorithm presented the lowest performance with  $R^2$  of 0.74, RMSE of 213.58 kg ha<sup>-1</sup> and EME of 15.06 kg ha<sup>-1</sup>. In a second study, we sought to estimate the volume of eucalyptus wood in the Brazilian cerrado using machine learning techniques and only climate data as model inputs, covering different growth ages in two periods of the year between January and June and between July and December. The models showed excellent results in estimating the volume of wood. The Random Forest model presented the best metrics during training and testing with  $R^2= 0.93$  and RMSE = 18.36 m<sup>3</sup>ha<sup>-1</sup> for the January-June model and  $R^2= 0.92$  and RMSE = 19.52 m<sup>3</sup>ha<sup>-1</sup> for the July-December model.

**Keywords:** agriculture, forest, agrometeorology, machine learning

**LISTA DE ABREVIATURAS E SIGLAS**

IA – Inteligência Artificial

ML – Machine learning

RNAs – Redes Neurais Artificiais

DNNs – Redes Neurais Profundas

CNNs- Redes Neurais Convolucionais

LSTMs - Long Short Term Memory (Memória de longo prazo)

MLP - Multilayer Perceptron

GEE – Gases do Efeito Estufa

RUE - Eficiência no uso da radiação

T – Temperatura do ar (°C)

P – Precipitação Pluviométrica (mm)

AWC – Capacidade de água disponível (mm)

CET – Evapotranspiração da cultura (mm)

AET – Evapotranspiração real (mm)

STO – Armazenamento de água no solo (mm)

DEF – Déficit Hídrico (mm)

EXC – Excedente Hídrico (mm)

BH – Balanço Hídrico

CV —Validação Cruzada

RF – Random Forest

SVM – Suport Vector Machine (Máquina de vetores de suporte)

XGBoost – Extreme Gradient Boosting

## LISTA DE FIGURAS

**Figura 1 (Capítulo 2)** - Mapa de localização do MATOPIBA e região Sul do Maranhão.

**Figura 2 (Capítulo 2)** - Representação climática da área de estudo: (A) temperatura média (°C) da série histórica (2008–2017); (B) média de precipitação (mm) da série histórica (2008–2017).

**Figura 3 (Capítulo 2)** - Representação dos dados de entrada nos modelos. Legenda: T – Temperatura (°C); P – Precipitação (mm); CET – Evapotranspiração da cultura (mm); AET – Evapotranspiração real da cultura (mm); STO – Armazenamento (mm); DEF – Déficit (mm); EXC – Excesso (mm)

**Figura 4 (Capítulo 2)** - Mapa térmico de correlação agrupado segundo clima e fases fenológicas da soja.

**Figura 5 (Capítulo 2)** - Modelos de aprendizado de máquina, Floresta Aleatória - RF, Redes Neurais Artificiais - RNA, Máquinas de Vetores de Suporte - kernel - Função de base radial - SVM\_RBF, Máquinas de Vetores de Suporte - kernel - Linear - SVM\_LINEAR e Máquinas de Vetores de Suporte - kernel - Polinômio de terceira ordem - SVM\_POLY, para previsão da produtividade da soja um mês antes da colheita.

**Figura 6 (Capítulo 2)** - Boxplot da produtividade observado e previsto por cada modelo de aprendizado de máquina. Legenda: losango = média, linha vermelha = mediana, caixa = 50% dos dados e traço = 99% dos dados.

**Figura 7 (Capítulo 2)** - Mapa da distribuição da produtividade de soja: A: produtividade observada, B: produtividade prevista por Random Forest, C: produtividade prevista por RNA, D: produtividade prevista por SVM\_RBF, E: produtividade prevista por SVM\_LINEAR, F: produtividade prevista por SVM\_POLY.

**Figura 8 (Capítulo 2)** - Mapa da diferença entre a produtividade prevista e observada para cada modelo.

**Figura 1 (Capítulo 3)** – Mapa de localização das áreas de estudo e descrição do bioma Cerrado. As unidades de produção foram divididas aleatoriamente em um conjunto de treino (em vermelho – 70% do total de dados) e um conjunto de teste (em amarelo – 30% do total de dados)

**Figura 2 (Capítulo 3)** - Distribuição dos valores de precipitação mensal para série histórica de 12 anos. O verão que ocorre entre dezembro e fevereiro é considerado chuvoso enquanto o inverno ocorrendo entre junho e agosto é caracterizado pelos baixos valores de precipitação na região do Cerrado – Mato Grosso do Sul, Brasil.

**Figura 3 (Capítulo 3)** - Temperatura média do ar mensal da área de estudo para o período de 12 anos com temperaturas na região do Cerrado – Mato Grosso do Sul, Brasil.

**Figura 4 (Capítulo 3)** - Evapotranspiração potencial mensal na região do Cerrado – Mato Grosso do Sul, Brasil.

**Figura 5 (Capítulo 3)** – Distribuição dos valores de armazenamento mensal de água no solo para série histórica de 12 anos na região do Cerrado – Mato Grosso do Sul, Brasil.

**Figura 6 (Capítulo 3)** - Distribuição dos valores mensais de deficiência de água no solo para série histórica de 12 anos na região do Cerrado – Mato Grosso do Sul, Brasil

**Figura 7 (Capítulo 3)** - Distribuição dos valores de excedente de água no solo para série histórica de 12 anos na região do Cerrado – Mato Grosso do Sul, Brasil.

**Figura 8 (Capítulo 3)** - Curva de avaliação do volume total com casca ( $m^3 ha^{-1}$ ) ao longo do crescimento do eucalipto com medições a partir dos 2 anos de idade até aos 7 anos. O volume total com casca é avaliado a partir do segundo ano, apresentando tendência de crescimento até aos 7 anos com incrementos anuais

**Figura 9 (Capítulo 3)** - Curva de avaliação do diâmetro a altura do peito (cm) ao longo do crescimento do eucalipto com medições a partir dos 2 anos de idade até aos 7 anos.

**Figura 10 (Capítulo 3)** - Curva de avaliação da altura dominante (m) ao longo do crescimento do eucalipto com medições a partir dos 2 anos de idade até aos 7 anos.

**Figura 11 (Capítulo 3)** - Correlação de Spearman das variáveis analisadas com a produtividade; a) período de janeiro a junho; b) período de julho a dezembro; VTCC – volume total com casca, P – precipitação, Tmax – temperatura máxima do ar, Tmin – temperatura mínima do ar, Tmean – temperatura média do ar, SR – radiação solar global, RH – umidade relativa, PET – Evapotranspiração potencial, STO – armazenamento de água no solo, DEF – Deficiência Hídrica, EXC – excedente hídrico.

**Figura 12 (Capítulo 3)** - Esquema de entrada de dados para os modelos sazonais de Random forest – RF, Regressão Linear Multipla – MLR, Suporte Vector Machine – SVM e XGBoost Regressor – XGB em que, P – precipitação (mm), Tmax – temperatura máxima do ar ( $^{\circ}C$ ), Tmin – temperatura mínima do ar ( $^{\circ}C$ ), Tmean – temperatura média do ar( $^{\circ}C$ ), RH – umidade relativa,  $Q_0$  – radiação no topo da atmosfera ( $W m^{-2}$ ), PET – evapotranspiração potencial (mm), STO – armazenamento de água no solo (mm), DEF deficiência hídrica (mm), EXC excedente hídrico (mm).

**Figura 13 (Capítulo 3)** – Seleção de variáveis via stepwise considerando ‘p-value’ menor que 0,05 como condição para determinação das variáveis a serem utilizadas para treinameto dos modelos.

**Figura 14 (Capítulo 3)** - Regressão Linear Multipla com Stepwise backward para seleção da variáveis; a) Ajuste inicial do modelo de regressão com 71 variáveis

independentes para o período de janeiro a junho; b) Resultado do modelo de regressão após stepwise backward, com 47 variáveis para o período de janeiro a junho; c) Ajuste inicial do modelo de regressão com 71 variáveis independentes para o período de julho a dezembro; d) Resultado do modelo de regressão após stepwise backward, com 45 variáveis para o período de julho a dezembro.

**Figura 15 (Capítulo 3)** - Teste dos modelos Random forest – RF, Regressão Linear Multipla – MLR, Suporte Vector Machine – SVM e XGBoost Regressor – XGB, com conjunto de dados climáticos entre janeiro e julho da série histórica de 12 anos, considerando florestas entre 2 e 7 anos de idade na região do Cerrado – Mato Grosso do Sul, Brasil.

**Figura 16 (Capítulo 3)** – Teste dos modelos Random forest – RF, Regressão Linear Multipla – MLR, Suporte Vector Machine – SVM e XGBoost Regressor – XGB, com conjunto de dados climáticos entre janeiro e julho da série histórica de 12 anos, considerando florestas entre 2 e 7 anos de idade na região do Cerrado – Mato Grosso do Sul, Brasil.

**Figura 17 (Capítulo 3)** - Volume estimado ao longo do crescimento do eucalipto, considerando os valores médios para cada idade do conjunto de teste dos modelos (janeiro-junho) Random forest – RF, Regressão Linear Multipla – MLR, Suporte Vector Machine – SVM e XGBoost Regressor – XGB, para a região do Cerrado – Mato Grosso do Sul, Brasil.

**Figura 18 (Capítulo 3)** - Valores médios de volume total com casca por idade utilizados no conjunto de teste dos modelos seguido dos resultados médios estimados com avaliação dos resíduos entre o estimado e observado para o modelo de janeiro-junho.

**Figura 19 (Capítulo 3)** – Comportamento da curva do volume estimado ao longo do crescimento do eucalipto, considerando os valores médios para cada idade do conjunto de teste dos modelos (julho-dezembro) Random forest – RF, Regressão Linear Multipla – MLR, Suporte Vector Machine – SVM e XGBoost Regressor – XGB, para a região do Cerrado – Mato Grosso do Sul, Brasil.

**Figura 20 (Capítulo 3)** - Valores médios de volume total com casca por idade utilizados no conjunto de teste dos modelos seguido dos resultados médios estimados com avaliação dos resíduos entre o estimado e observado para o modelo de julho-dezembro.

## **CAPÍTULO 1 – Considerações gerais**

### **1.1 Introdução**

O Cerrado brasileiro é o segundo maior bioma da América do Sul, cobrindo originalmente aproximadamente 2 milhões de km<sup>2</sup> (Glatzle et al., 2024). É um ecossistema heterogêneo formado por um mosaico de ambientes abertos e fisionomias de vegetação fechada, composta por pastagens, savanas e florestas estacionais (Costa et al., 2020). No entanto, este bioma passou por importantes transformações nas últimas cinco décadas (Hunke et al., 2015).

O avanço antrópico sobre o bioma Cerrado está intrinsecamente ligado ao desenvolvimento da região a qual ele ocupa no território nacional (Ruas et al., 2022). O Cerrado é considerado uma grande região biogeográfica que possui riqueza sociocultural e ecológica inestimável, apresenta uma rica biodiversidade e um mosaico de fitofisionomias que se aglutinam sob relações de interdependência entre clima, solo, relevo e vegetação (de Souza et al., 2023).

A expansão da agricultura e o uso de tecnologias no Cerrado geraram benefícios socioeconômicos inegáveis: aumento da oferta dos produtos agrícolas, ganhos na produtividade da agricultura, diversificação das economias locais e aumento da renda de municípios, e melhorias sociais em várias localidades (Klink et al., 2005).

As transformações ocorridas no Cerrado trouxeram grandes desafios para o bioma que sofreu com a fragmentação de habitats, mitigação da biodiversidade, degradação de ecossistemas, desequilíbrios no ciclo do carbono e possível modificações climáticas regionais (Klink et al., 2005).

Ao mesmo tempo, esse avanço foi parte essencial do crescimento da economia brasileira, que até a década de 1960 ainda importava alimentos, como milho, arroz, cereais e carne de frango. Atualmente o país é um ativo exportador de alimentos, principalmente carnes e grãos (Carneiro Filho et al., 2016), com destaque para a soja produzida no Cerrado, ocupando a primeira posição no ranking exportador global e o segundo em produção cujo avanços tecnológicos foram o principal motivo que tornou isso possível (Magalhães et al., 2020).

O Cerrado brasileiro compreende 1297 municípios com diferentes históricos de ocupação em 12 estados brasileiros Bahia, Distrito Federal, Goiás,

### **3.4. Conclusões**

Todos os modelos apresentam elevadas acurácias, porém o Random Forest caracterizou-se como o melhor modelo, apresentando as melhores métricas durante o treinamento e teste com,  $R^2= 0.93$  e  $RMSE = 18.36 \text{ m}^3\text{ha}^{-1}$  para o modelo janeiro-junho e  $R^2= 0.92$  e  $RMSE = 19.52 \text{ m}^3\text{ha}^{-1}$  para o modelo de julho-dezembro.

Para o período de janeiro a junho, o modelo Suporte Vector machine - SVM e a Regressão Linear Multipla – RLM tendem a superestimar os resultados em  $2 \text{ m}^3\text{ha}^{-1}$  e de  $14 \text{ m}^3\text{ha}^{-1}$  para as idades de 6 e 7 anos respectivamente, enquanto os modelos Random Forest e o XGB subestimam os resultados nessas idades.

Para o período de julho a dezembro, os modelos Random Forest superestimam os resultados em todas as idades, enquanto o modelo SVM subestimam valores nas idades 2, 5, 6 e 7 anos, com diferenças de  $-1 \text{ m}^3\text{ha}^{-1}$ ,  $-2 \text{ m}^3\text{ha}^{-1}$ ,  $-8 \text{ m}^3\text{ha}^{-1}$  e  $-2 \text{ m}^3\text{ha}^{-1}$  respectivamente, em relação ao valor observado.

## REFERÊNCIAS

Alvares, C. A., Stape, J. L., Sentelhas, P. C., Gonçalves, J. D. M., & Sparovek, G. (2013). Köppen's climate classification map for Brazil. **Meteorologische zeitschrift**, 22(6), 711-728.

Alvares, Clayton Alcarde et al. Decision-Tree Application to Predict and Spatialize the Wood Productivity Probabilities of Eucalyptus Plantations. **Forests**, v. 14, n. 7, p. 1334, 2023.

Billings, W. Dwight. The environmental complex in relation to plant growth and distribution. **The Quarterly Review of Biology**, v. 27, n. 3, p. 251-265, 1952.

Breiman, Leo. Random forests. **Machine learning**, v. 45, p. 5-32, 2001.

Campoe, Otávio C. et al. Meteorological seasonality affecting individual tree growth in forest plantations in Brazil. **Forest Ecology and Management**, v. 380, p. 149-160, 2016.

Chen, Tianqi; GUESTRIN, Carlos. Xgboost: A scalable tree boosting system. In: **Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining**. 2016. p. 785-794.

Christina, Mathias et al. Importance of deep water uptake in tropical eucalypt forest. **Functional Ecology**, v. 31, n. 2, p. 509-519, 2017.

Cordeiro, Márcio Assis et al. Volumetric estimates in eucalyptus stands using support vector machines and artificial neural networks. **Madera y bosques**, v. 28, n. 1, 2022.

de Alcântara, Aline Edwiges Mazon et al. Use of artificial neural networks to assess yield projection and average production of eucalyptus stands. **African Journal of Agricultural Research**, v. 13, n. 42, p. 2285-2297, 2018.

de Oliveira Neto, Ricardo Rodrigues et al. Estimation of Eucalyptus productivity using efficient artificial neural network. **European Journal of Forest Research**, v. 141, n. 1, p. 129-151, 2022.

de Souza Diniz, Juliana Maria Ferreira et al. Estimating stem volume of Eucalyptus sp. and Pinus sp. plantations in Brazil, using Sentinel-1B and ALOS-2/PALSAR-2 data. **Journal of Applied Remote Sensing**, v. 17, n. 1, p. 014513-014513, 2023.

Diaz, Gonzalo I. et al. An effective algorithm for hyperparameter optimization of neural networks. **IBM Journal of Research and Development**, v. 61, n. 4/5, p. 9: 1-9: 11, 2017.

Elli, Elvis Felipe et al. Intercomparison of structural features and performance of Eucalyptus simulation models and their ensemble for yield estimations. **Forest Ecology and Management**, v. 450, p. 117493, 2019.

Elli, Elvis Felipe; SENTELHAS, Paulo Cesar; BENDER, Fabiani Denise. Impacts and uncertainties of climate change projections on Eucalyptus plantations productivity across Brazil. **Forest Ecology and Management**, v. 474, p. 118365, 2020.

Elshawi, Radwa; MAHER, Mohamed; SAKR, Sherif. Automated machine learning: State-of-the-art and open challenges. **arXiv preprint arXiv: 1906.02287**, 2019.

FAO 2020 Global Forest Resources Assessment 2020 (Rome: FAO) (available at: <https://doi.org/10.4060/ca8753en>)

He, Haibo et al. Incremental learning from stream data. **IEEE Transactions on Neural Networks**, v. 22, n. 12, p. 1901-1914, 2011.

Li, Ying chang et al. Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms. **Scientific reports**, v. 10, n. 1, p. 9952, 2020.

Liu, Yingxia et al. Analysis of spatio-temporal variation of crop yield in China using stepwise multiple linear regression. **Field Crops Research**, v. 264, p. 108098, 2021.

Maulud, Dastan; ABDULAZEEZ, Adnan M. A review on linear regression comprehensive in machine learning. **Journal of Applied Science and Technology Trends**, v. 1, n. 4, p. 140-147, 2020.

Messier, C., Bauhus, J., Sousa-Silva, R., Auge, H., Baeten, L., Barsoum, N., ... & Zemp, D. C. (2022). For the sake of resilience and multifunctionality, let's diversify planted forests!. **Conservation Letters**, 15(1), e12829.

Otto, Marina Shinkai Gentil et al. Photosynthesis, stomatal conductance and productivity of Eucalyptus clones under different soil and climatic conditions. **Revista Árvore**, v. 37, p. 431-439, 2013.

Pedregosa, Fabian et al. Scikit-learn: Machine learning in Python. **the Journal of machine Learning research**, v. 12, p. 2825-2830, 2011.

Ramantswana, Muedanyi; GUERRA, Saulo Philipe Sebastião; ERSSON, Back Tomas. Advances in the mechanization of regenerating plantation forests: A review. **Current forestry reports**, v. 6, p. 143-158, 2020.

Rowland, Lucy et al. The sensitivity of wood production to seasonal and interannual variations in climate in a lowland Amazonian rainforest. **Oecologia**, v. 174, p. 295-306, 2014.

Santana, Dthenifer Cordeiro et al. Machine Learning Methods for Woody Volume Prediction in Eucalyptus. **Sustainability**, v. 15, n. 14, p. 10968, 2023.

Scolforo, Henrique Ferraco et al. Eucalyptus growth and yield system: Linking individual-tree and stand-level growth models in clonal Eucalypt plantations in Brazil. **Forest Ecology and Management**, v. 432, p. 1-16, 2019.

Seng Hua, Lee et al. Engineering wood products from Eucalyptus spp. **Advances in Materials Science and Engineering**, v. 2022, p. 1-14, 2022.

Shaharum, N. S. N. et al. Image classification for mapping oil palm distribution via support vector machine using scikit-learn module. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 42, p. 133-137, 2018.

Silva, Jeferson Pereira Martins et al. Prognosis of forest production using machine learning techniques. **Information Processing in Agriculture**, 2021.

Silva, L. C. R. Dinâmica de transição e interações entre fitofisionomias florestais e formações vegetacionais abertas do bioma Cerrado. Brasília-DF. 2007. 168p. Dissertação (Mestrado em Ciências Florestais) – Departamento de Engenharia Florestal, Universidade de Brasília, Brasília, 2007.

Stape, Jose Luiz et al. The Brazil Eucalyptus Potential Productivity Project: Influence of water, nutrients and stand uniformity on wood production. **Forest Ecology and Management**, v. 259, n. 9, p. 1684-1694, 2010.

Stape, Jose Luiz et al. Water use, water limitation, and water use efficiency in a Eucalyptus plantation. **Revista Bosque**, v. 25, n. 2, p. 35-41, 2004.

Tehrany, Mahyat Shafapour et al. Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. **Catena**, v. 125, p. 91-101, 2015.

Thornton, P. K., Ericksen, P. J., Herrero, M., & Challinor, A. J. (2014). Climate variability and vulnerability to climate change: a review. **Global change biology**, 20(11), 3313-3328.

Van de Ven, Gido M.; TUYTELAARS, Tinne; TOLIAS, Andreas S. Three types of incremental learning. **Nature Machine Intelligence**, v. 4, n. 12, p. 1185-1197, 2022.

Vieira, Giovanni Correia et al. Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. **Science of the Total Environment**, v. 619, p. 1473-1481, 2018.

Yang, Li; SHAMI, Abdallah. On hyperparameter optimization of machine learning algorithms: Theory and practice. **Neurocomputing**, v. 415, p. 295-316, 2020.

Zhou, Xiaoguo et al. Effects of understory management on trade-offs and synergies between biomass carbon stock, plant diversity and timber production in eucalyptus plantations. **Forest Ecology and Management**, v. 410, p. 164-173, 2018.