



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM DESENVOLVIMENTO E MEIO
AMBIENTE

FERNANDO ANTONIO MARÇAL GARCIA

**ADSORÇÃO DE PESTICIDA POR BIOCARVÃO: modelagem com variáveis
edafoclimáticas por ensemble RNA**

Recife
2025

FERNANDO ANTONIO MARÇAL GARCIA

**ADSORÇÃO DE PESTICIDA POR BIOCARVÃO: modelagem com variáveis
edafoclimáticas por ensemble RNA**

Tese apresentada ao Programa de Pós-Graduação em Desenvolvimento e Meio Ambiente da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Doutor em Desenvolvimento e Meio Ambiente.

Área de concentração: Tecnologias para o Desenvolvimento Sustentável.

Orientadora: Prof^a. Dr^a. Maria do Socorro Bezerra de Araújo

Coorientador: Prof. Dr. André Maciel Netto

Recife

2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Garcia, Fernando Antonio Marçal.

Adsorção de pesticida por biocarvão: modelagem com variáveis edafoclimáticas por ensemble RNA / Fernando Antonio Marçal Garcia. - Recife, 2025.
192f.: il.

Tese (Doutorado) - Universidade Federal de Pernambuco, Centro de Filosofia e Ciências Humanas, Programa de Pós-Graduação em Desenvolvimento e Meio Ambiente, 2025.

Orientação: Maria do Socorro Bezerra de Araújo.

Coorientação: André Maciel Netto.

Inclui referências e apêndices.

1. Agricultura de precisão; 2. Otimização de processos; 3. Inteligência Artificial; 4. Evapotranspiração; 5. Condutividade Hidráulica Saturada. I. Araújo, Maria do Socorro Bezerra de. II. Maciel Netto, André. III. Título.

UFPE-Biblioteca Central

FERNANDO ANTONIO MARÇAL GARCIA

ADSORÇÃO DE PESTICIDA POR BIOCARVÃO: modelagem com variáveis edafoclimáticas por ensemble RNA

Tese apresentada ao Programa de Pós-Graduação em Desenvolvimento e Meio Ambiente da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Doutor em Desenvolvimento e Meio Ambiente.
Área de concentração: Tecnologias para o Desenvolvimento Sustentável

Aprovada em: 00/02/2025

BANCA EXAMINADORA

Prof^a. Dr^a. Maria do Socorro Bezerra de Araújo (Orientadora)
Universidade Federal de Pernambuco – UFPE

Prof^a. Dr^a. Josiclêda Domiciano Galvíncio (Examinadora Interna)
Universidade Federal de Pernambuco – UFPE

Prof. Dr. Raimundo Aprígio de Menezes Junior (Examinador Externo)
Universidade Federal de Pernambuco – UFPB

Prof. Dr. Cleber Zanchettin (Examinador Externo)
Universidade Federal de Pernambuco – UFPE

Prof. Dr. Antonio Cláudio Marques Afonso (Examinador Externo)
– UF.

Prof. Dr. Fernando Cartaxo Rolim Neto (Examinador Externo)
Universidade Federal Rural de Pernambuco - UFRPE

SUPLENTE:

Profa. Dra. Jarcilene Silva de Almeida (Examinador Suplente Interno)

Universidade Federal de Pernambuco – UFPE

Profa. Dra. Jéssyka Flavianne Ferreira Vilela (Examinador Suplente Externo)

Universidade Federal de Pernambuco – UFPE

Aos meus filhos, Renata e Rodolfo e minha esposa Fátima,
que lograram grandes sacrifícios, para eu chegar aqui.

Dedico

AGRADECIMENTOS

Agradeço a todos que, de forma direta ou indireta, ajudaram-me a chegar até aqui, particularmente, aos meus orientadores, Prof^a. Dr^a. Socorro e Prof. Dr. André, Prof^a. Dr^a. Jarcilene Almeida, colegas de doutorado, professores e servidores do PRODEMA.

Obrigado, Pai, por tudo que me destes e
Por tudo que o sou.
Obrigado, pelo dia de hoje,
Por tudo de bom que me ocorreu,
Por todos os desafios que me permitistes enfrentar,
Por tudo o que consegui com a Vossa graça ... (Livro dos Salmos, 2011).

“Tenho aprendido a oferecer menos resistência aos sacrifícios impostos, a suportar melhor as dificuldades, a ser mais tolerante. E a encontrar um sentido real para a vida. Apesar dos que a tudo isso se opõem [...]”. (Em “Apesar de Vocês”, Tom Coelho).

RESUMO

A complexidade e extensão de áreas de atuação das ciências ambientais abrange, principalmente, o planejamento ambiental, o manejo, conservação, gestão e monitoramento de recursos, como uso racional da biodiversidade e preservação responsável da mobilidade urbana e social, caracterizando sua multidisciplinariedade. Desenvolver pesquisas sobre processos ambientais, muitos deles fortemente correlacionados, implica no delineamento de variáveis complexas, com alta variedade espacial e distribuição não linear, tais como micro e macro nutrientes no meio, índices de temperatura, pressão e evapotranspiração, umidade climática, balanço de energia, além das decorrentes de ação de agentes externos, de origem físico-química e biológicas, estranhas ao bioma estudado. A modelagem dos diversos experimentos ambientais por técnicas estatísticas, matemáticas e computacionais tem auxiliado na compreensão mais abrangente, precisa e preditiva de seus impactos. A rápida evolução da inteligência artificial (IA) tem afetado positivamente técnicas e ferramentas utilizadas para este fim. Uma grande vantagem da aplicação de modelagem de bioprocessos com IA, se dá na replicação de experimentos realizados com investigação de novos parâmetros, variáveis, ou ampliação de cenários, na busca de predição e otimização de resultados. Nos processos de remediação ou melhoramento de solos por poluentes diversos, o biocarvão tem sido investigado, como proposta alternativa a produtos sintéticos, principalmente de origem agroquímica. Os desafios para sua disseminação e ampla utilização passam por minorar sua complexidade fabril, compreensão de seus mecanismos de atuação, otimização de seus resultados. Outros aspectos investigados relatam preocupação com ampliação do espectro funcional, bem como de seu custo operacional em escala. Assim, este trabalho revisou a aplicação prática dos modelos de IA em sistemas de gerenciamento da atividade agrícola, como sistemas de agricultura de precisão (AP 4.0), indicando lacunas a serem desenvolvidas, bem como os avanços esperados, com a modernização pelos avanços trazidos pela evolução de técnicas agrícolas com auxílio de IA. Outro importante aspecto investigado foi avaliar como a introdução de duas variáveis edafoclimáticas, evapotranspiração real (ET_r) e condutividade hidráulica saturada (K_s), pode contribuir na formulação, compreensão e melhoria de resultados de adsorção um biocarvão, quando aplicado a um solo contendo concentrações do herbicida Imazapic®. Esta avaliação foi realizada mediante a

predição de resultados obtidos por modelagem computacional, utilizando poderosos e recentes algoritmos de inteligência artificial. Os resultados alcançados na modelagem do processo produtivo do biocarvão com introdução destas duas variáveis, mostraram que a capacidade adsortiva do biocarvão foi otimizada. Além disso, propiciou melhor compreensão da influência das variáveis envolvidas no processo, possibilitando ampliação de futuras investigações. Este trabalho inova ao demonstrar a viabilidade técnica da modelagem de processos produtivos de biocarvão serem flexibilizados, ampliando-se e personalizando seu espectro de atuação, através das alterações dinâmicas de seus atributos-chave, permutação de biomassa, ou adição de variáveis complexas, intrínsecas ao seu objetivo de aplicação, com melhor compreensão de seus mecanismos construtivos em meios altamente complexos, como solos afetados por poluentes multissítios, cada vez mais comuns, com bons resultados, ressaltando o papel e contribuição das ferramentas computacionais de inteligência artificial como suporte e auxílio aos processos envolvidos, contribuindo assim, para o avanço das pesquisas ambientais.

Palavras-chave: Agricultura de precisão; otimização de processos; inteligência artificial; evapotranspiração; condutividade hidráulica saturada.

ABSTRACT

The complexity and breadth of the areas of activity of environmental sciences mainly encompass environmental planning, management, conservation, management and monitoring of resources, such as the rational use of biodiversity and responsible preservation of urban and social mobility, characterizing its multidisciplinary. Developing research on environmental processes, many of which are strongly correlated, involves outlining complex variables, with high spatial variety and non-linear distribution, such as micro and macro nutrients in the environment, temperature, pressure and evapotranspiration indices, climatic humidity, energy balance, in addition to those resulting from the action of external agents, of physical-chemical and biological origin, foreign to the biome studied. Modeling the various environmental experiments using statistical, mathematical and computational techniques has helped in a more comprehensive, accurate and predictive understanding of their impacts. The rapid evolution of artificial intelligence (AI) has positively affected the techniques and tools used for this purpose. A great advantage of applying bioprocess modeling with AI is the replication of experiments carried out with the investigation of new parameters, variables, or expansion of scenarios, in the search for prediction and optimization of results. In soil remediation or improvement processes for various pollutants, biochar has been investigated as an alternative proposal to synthetic products, mainly of agrochemical origin. The challenges for its dissemination and widespread use include reducing its manufacturing complexity, understanding its operating mechanisms, and optimizing its results. Other aspects investigated report concerns about expanding the functional spectrum, as well as its operational cost on a large scale. Thus, this work reviewed the practical application of AI models in agricultural activity management systems, such as precision agriculture systems (AP 4.0), indicating gaps to be developed, as well as the expected advances, with the modernization by the advances brought by the evolution of agricultural techniques with the aid of AI. Another important aspect investigated was to evaluate how the introduction of two edaphoclimatic variables, real evapotranspiration (ET_r) and saturated hydraulic conductivity (K_s), can contribute to the formulation, understanding and improvement of biochar adsorption results, when applied to a soil containing concentrations of the herbicide Imazapic®. This evaluation was carried out by predicting results obtained by computational modeling, using powerful and recent artificial intelligence algorithms. The results

achieved in the modeling of the biochar production process with the introduction of these two variables showed that the adsorptive capacity of the biochar was optimized. Furthermore, it provided a better understanding of the influence of the variables involved in the process, enabling the expansion of future investigations. This work innovates by demonstrating the technical feasibility of modeling biochar production processes to be made more flexible, expanding and personalizing their spectrum of action, through dynamic changes in their key attributes, biomass permutation, or addition of complex variables, intrinsic to their application objective, with a better understanding of their constructive mechanisms in highly complex environments, such as soils affected by increasingly common multi-site pollutants, with good results, highlighting the role and contribution of artificial intelligence computational tools as support and assistance to the processes involved, thus contributing to the advancement of environmental research.

Keywords: Precision agriculture; process optimization; artificial intelligence; evapotranspiration; saturated hydraulic conductivity.

LISTA DE ILUSTRAÇÕES

TESE

Figura 1 –	Evolução das pesquisas com biocarvão	23
Figura 2 –	Evolução das pesquisas com biocarvão	24
Figura 3 –	Evolução das pesquisas com biocarvão	28
CAPÍTULO 1 – PREDIÇÃO DE ÍNDICE DE EVAPOTRANSPIRAÇÃO REAL DIÁRIA AVALIADA POR MODELOS DE INTELIGÊNCIA ARTIFICIAL		28
Descrição das etapas aplicadas para avaliação dos modelos de DL e ML para a estimação da evapotranspiração real diária (ETr)		33
Figura 1 –	Localização geográfica da área de estudo, São João/PE	35
Figura 3 –	Box plot dos pontos extremos por variável, no conjunto amostral original	36
Figura 4 –	Gráficos de Resíduos e Dispersão da variável-alvo (ETr)	43
Figura 5 –	Gráfico de Curvas de Perdas durante o treinamento do modelo DenseNet	44
Figura 6 –	Gráficos de regularização L1 e L2 durante processamento dos modelos de IA	45
Figura 7 –	Gráfico de análise de dispersão entre valores de MSE e R ² para todos os modelos de IA avaliados	46
Figura 8 –	Gráfico comparativo de métricas normalizadas de desempenho entre modelos	47
Figura 9 –	Gráficos comparativo complementares de desempenho do modelo DenseNet	48
Figura 10 –	Desempenho da DenseNet das fases de treinamento e teste	49
Figura 11 –	Análise da forma de relacionamento entre variáveis preditoras e alvo	51
Figura 12 –	Gráficos SHAP com direção, complexidade e relacionamento entre NDVI e ETr	52
Figura 13 –	Gráfico de correlação entre as variáveis preditoras e alvo do modelo de IA.	54
Figura 14 –	Mapa interativo da modelagem da DenseNet, comparativamente aos dados reais.	55
CAPÍTULO 2 – PREDIÇÃO DA CONDUTIVIDADE HIDRÁULICA SATURADA AVALIADA POR MODELOS DE INTELIGÊNCIA ARTIFICIAL		61
Figura 1 –	Descrição das etapas aplicadas para avaliação dos modelos de ML para a estimação da condutividade hidráulica saturada (Ks)	66

Figura 2 –	Localização geográfica de São João/PE	67
Figura 3 –	Box plot dos pontos extremos por variável, no conjunto amostral original.....	69
Figura 4 –	Gráficos de desempenho (valores reais versus previstos) de Ks pelos modelos de IA	74
Figura 5 –	Gráfico de geração de dados sintéticos	79
Figura 6 –	Visualização das variáveis de estudo nos limites superior e inferior (scatter plots), caracterizando pontos extremos ou outliers do conjunto amostral experimental original.....	82
Figura 7 –	Visualização gráfica QQ-plots das variáveis de estudo nos conjuntos de dados originais e expandidos.....	87
Figura 8 –	Visualização das variáveis de estudo nos limites superior e inferior (scatter plots), caracterizando pontos extremos ou outliers dos conjuntos sintéticos expandidos com e sem outliers.....	89
Figura 9 –	Gráfico de análise de dispersão entre valores de MSE e R2 para todos os modelos de ML avaliados	91
Figura 10 –	Gráficos de comparativo entre as métricas normalizadas e de evolução das métricas de desempenho entre modelos para avaliação de desempenho dos modelos de ML.....	92
Figura 11 –	Gráficos de regularização L1 e L2 e PCA de correlação da variável-alvo (Ks), durante processamento dos modelos de ML.....	94
Figura 12 –	Desempenho da RNA na estimativa de Ks, avaliado pelo número de neurônios na camada escondida.....	95
Figura 13 –	Análise da forma de relacionamento entre variáveis preditoras e alvo.....	98
Figura 14 –	Gráficos SHAP com direção, complexidade e relacionamento entre variáveis na modelagem IA.....	101
Figura 15 –	Gráfico de correlação entre as variáveis e seus impactos sobre o modelo de ML.....	104
Figura 16 –	Contribuição das variáveis preditoras para a variabilidade da modelagem IA	105
Figura 17 –	Simulação da equação de FPT com dados de Ks preditos pelo modelo da RNA.....	107
CAPÍTULO 3 – INFLUÊNCIA DE VARIÁVEIS EDAFOCLIMÁTICAS NA AVALIAÇÃO ADSORTIVA DE UM BIOCÁRVÃO POR MODELO ENSEMBLE DE INTELIGÊNCIA ARTIFICIAL		114
Figura 1 –	Localização geográfica da área de estudo (coleta de ETr e Ks), em São João/PE	117

Figura 2 –	Esquema metodológico de desenvolvimento da pesquisa.....	119
Figura 3 –	Distribuição espacial dos dados sintéticos – a) variável C_{eq} ; b) variável S	120
Figura 4 –	Densidade espacial dos dados sintéticos	122
Figura 5 –	Gráficos dos dados sintéticos – a) resíduos; b) dispersão.....	124
Figura 6 –	Gráfico de Curvas de Perdas durante o treinamento do modelo DenseNet.....	125
Figura 7 –	Gráficos dos modelos IA avaliados – a) desempenho; b) acurácia MSE vs R^2	126
Figura 8 –	Gráficos dos modelos IA avaliados – a) acurácia MSE vs Neurônios; b) topologia	127
Figura 9 –	Desempenho da DenseNet nas fases de treinamento e teste – a) C_{eq} ; b) S; c) ETr; d) K_s	128
Figura 10 –	Análise SHAP (direção, complexidade e relacionamento) entre C_{eq} e S estimadas.....	130
Figura 11 –	Análises explicativas da estimação – a) Contribuições Integrated Gradients; b) LIME.....	131
Figura 12 –	Análises explicativas da estimação – a) Importância por permutação; b) Matriz de correlação.....	133
Figura 13 –	Análises comparativas entre as isothermas experimentais e estimadas – a) Freundlich originais e estimadas; b) Langmuir originais e estimadas; c) Freundlich vs Langmuir experimentais; Freundlich vs Langmuir estimadas.....	135

LISTA DE TABELAS

TESE

CAPÍTULO 1 – PREDIÇÃO DE ÍNDICE DE EVAPOTRANSPIRAÇÃO REAL DIÁRIA AVALIADA POR MODELOS DE INTELIGÊNCIA ARTIFICIAL	28
Tabela 1 – Análise descritiva do conjunto amostral original e estimado	38
Tabela 2 – Classificação por desempenho dos modelos de IA treinados com conjunto de dados sintéticos	41
CAPÍTULO 2 – PREDIÇÃO DA CONDUTIVIDADE HIDRÁULICA SATURADA AVALIADA POR MODELOS DE INTELIGÊNCIA ARTIFICIAL	61
Tabela 1 – Análise descritiva do conjunto amostral original.....	70
Tabela 2 – Análise descritiva do conjunto amostral sem poda de pontos extremos, expandido por GAN.....	72
Tabela 3 – Classificação por desempenho dos modelos de IA treinados com conjunto de dados sintéticos	74
Tabela 4 – Análise descritiva do conjunto de dados sintéticos sem poda de pontos extremos, expandido por GAN	85
Tabela 5 – Teste de Normalidade para o conjunto de dados originais sem poda de outliers	86
Tabela 6 – Teste de Normalidade para o conjunto de dados expandidos por GAN sem poda de outliers.	86
CAPÍTULO 3 – INFLUÊNCIA DE VARIÁVEIS EDAFOCLIMÁTICAS NA AVALIAÇÃO ADSORTIVA DE UM BIOCARVÃO POR MODELO ENSEMBLE DE INTELIGÊNCIA ARTIFICIAL	114
Tabela 1 – Resultados das métricas de acurácia para os quatro modelos de Inteligência Artificial avaliados	123
Tabela 2 – Valores de AIC e BIC baseados na estimação da DenseNet.....	134

LISTA DE ABREVIATURAS E SIGLAS

ETr	Evapotranspiração Real Diária
LSMs	Modelos de Superfície Terrestre
SR	Sensoriamento Remoto
SEBAL	Algoritmo Surface Energy Balance Algorithm for Land.
NDVI	Índice de Vegetação de Diferença Normalizada
IA	Inteligência Artificial
EEFLUX	Energy Balance for Forests at the Local Scale
DL	Aprendizado de Máquina Profundo
MOD16A2	Diagrama da topografia selecionada de RNA
ML	Aprendizado de Máquina
NASA	Agência de Administração Nacional da Aeronáutica e Espaço EUA
As	Classificação de solo no Sistema Brasileiro de Classificação de Solos
GEE	Google Earth Engine
IQR	Técnica de Interquartil
DENSENET	Densely Connected Convolutional Network
FCN	Fully Convolutional Network
U-NET	Convolutional Network for Biomedical Image Segmentation
MSE	Métrica de acurácia "Mean Squared Error"
RMSE	Métrica de acurácia "Root Mean Squared Error"
R ²	Métrica de acurácia "Coefficient of Determination"
MAE	Métrica de acurácia "Mean Absolute Error"
CC	Métrica de acurácia "Correlation Coefficient"
RF	Algoritmo de Aprendizado de Máquina Random Forest
L1/L2	Técnica de Regularização de Pesos do Modelo Neural
PDP	Partial Dependence Plots
SHAP	Shapley Additive ExPlanation
LIME	Técnica de Explicação Local
ALE	Técnica de Importância de Features por Permutação
Ceq	Concentração do Sóluto após Equilíbrio na Fase Líquida (mg/L)
S	Capacidade de Adsorção (mg/kg)
Ks	Condutividade Hidráulica Saturada
RNA	Rede Neural Artificial
DNN	Rede Neural Profunda
TPOT	Conjunto de algoritmos genéticos da biblioteca Python
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
IG	Técnica de Contribuições Integrated Gradients
GAN	Algoritmo de Inteligência Artificial
WGAN	Algoritmo de Inteligência Artificial
VAE	Algoritmo de Inteligência Artificial Variational Autoencoder
FPT	Funções de Pedotransferência

SUMÁRIO

1	INTRODUÇÃO	18
2	OBJETIVOS	20
2.1	OBJETIVO GERAL	20
2.2	OBJETIVOS ESPECÍFICOS	20
2.3	ORGANIZAÇÃO DO TRABALHO	20
3	REFERENCIAL TEÓRICO	22
3.1	BIOCARVÕES	22
3.2	MODELAGEM DE PROCESSOS AMBIENTAIS	25
4	CAPÍTULO 1 – PREDIÇÃO DE ÍNDICE DE EVAPOTRANSPIRAÇÃO REAL DIÁRIA AVALIADA POR MODELOS DE INTELIGÊNCIA ARTIFICIAL	28
5	CAPÍTULO 2 – PREDIÇÃO DA CONDUTIVIDADE HIDRÁULICA SATURADA AVALIADA POR MODELOS DE INTELIGÊNCIA ARTIFICIAL	61
6	CAPÍTULO 3 – INFLUÊNCIA DE VARIÁVEIS EDAFOCLIMÁTICAS NA AVALIAÇÃO ADSORTIVA DE UM BIOCARVÃO POR MODELO ENSEMBLE DE INTELIGÊNCIA ARTIFICIAL	114
7	CONCLUSÕES GERAIS	141
	REFERÊNCIAS	142
	APÊNDICE A – Artigo Publicado	151

1 INTRODUÇÃO

A disponibilidade global de alimentos frente à demanda populacional continua sendo um fator crítico (FAO-STAT, 2023), agravado principalmente nos países pobres ou em desenvolvimento, onde a ocorrência da concentração de renda é um catalisador dessa disparidade. Apesar das ações locais de governos, ou em nível global, por políticas promovidas pela ONU e outros organismos internacionais, a solução ainda passa pelo aumento da produção agrícola em termos mundiais, carreando diversos problemas associados, como a expansão sobre biomas nativos, aumento da poluição de ecossistemas, além dos custos operacionais por perdas. A atividade agrícola vem rapidamente se transformando da agricultura de subsistência para um modelo de negócios mundial (agronegócio), implementando mecanismos mais robustos de gestão de negócios (KIM, 2023), monitoramento, otimização de recursos (CHINTAKUNTA et al., 2023; TALAAT, 2023), como insumos, custos, previsões de produção, impactos ambientais, focando em aumento da competitividade e produtividade por hectare (EZZIYYANI et al., 2023; MAURYA et al., 2024), apoiados por modernas tecnologias industriais (BAZARGANI; DEEMYAD, 2024) e computacionais (LI et al., 2023; SARANYA et al., 2023; KRISHNABABU et al., 2024).

Em grande parte, os avanços tecnológicos incorporados no agronegócio voltado para a produção de alimentos (AASHU et al., 2024; SATHYA et al., 2024), está sendo realizado pelos sistemas aplicativos chamados agricultura de precisão – AP (GAWANDE et al., 2023, SANYAOLU; SADOWSKI, 2024; GOVINDARAJ et al., 2024; AASHU et al., 2024), ainda concentrado no segmento de médios e grandes produtores agrícolas, pouco acessível aos pequenos produtores (MIZIK, 2023). A solução de AP integra diversos módulos especializados para cada fase do processo produtivo agrícola, incorporando grandes inovações tecnológicas. Atualmente, a maioria das soluções de AP incorpora técnicas e procedimentos voltados para a preocupação ambiental, causadas por poluição residual do uso de agrotóxicos no agronegócio, como sua aplicação dosada e racional. Mais recentemente, as soluções de AP têm prospectado a viabilidade de se implementar produtos alternativos ao uso intensivo de pesticidas, como o biocarvão, a partir de estudos ofertados na literatura (SIM et al., 2021; KERNER et al., 2023; MAMATHA et al., 2024).

Essa perspectiva pode ser suportada pela modelagem inteligente do biocarvão, quanto a seu processo produtivo, destinação ou efeitos direcionados pela aplicação de AP. Compreender os mecanismos de produção do biocarvão, sua aplicação no meio, além dos resultados alcançados, é fator-chave para que sistemas de AP adotem o biocarvão

adequadamente. Assim, aprimorar pesquisas de modelagem de processos de biocarvão baseadas em ferramentas inteligência artificial, como aprendizado de máquina (CAI et al., 2023; CHINTAKUNTA et al., 2023; IJAZ; WOZNIAK, 2024; KUMAR et al., 2024; RAI; KUNTE, 2024) tem sido recentemente apresentado na literatura acadêmica. Desta forma, esta pesquisa investigou a aplicação prática de algoritmos de inteligência artificial, como alternativa viável e poderosa, impulsionando ampliação de novos estudos com o biocarvão, promovendo o aprofundamento na compreensão de seus mecanismos de formulação e atuação. A pesquisa também demonstra o potencial do uso do biocarvão em sistemas aplicados à agricultura de precisão e outros propósitos, contribuindo significativamente para o avanço das pesquisas em ciência ambiental.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Investigar a viabilidade qualitativa de um processo adsorptivo de pesticida Imazapic por um biocarvão, com inclusão de variáveis climáticas e do solo, por meio de modelagem por aplicação de algoritmos de aprendizado de máquina profundo (DL), utilizando dados experimentais.

2.2 OBJETIVOS ESPECÍFICOS

- a) Desenvolver dois algoritmos de aprendizado de máquina para, respectivamente, extrair dados climáticos de evapotranspiração real diária a partir de fonte satelital, assim como, estimar a condutividade hidráulica saturada do solo, a partir de fonte de dados experimentais;
- b) Desenvolver um modelo ensemble (modelo composto por um conjunto), integrando dois algoritmos de aprendizado de máquina a um modelo de rede neural profunda, para avaliar a capacidade adsorptiva de um biocarvão, com adição duas variáveis edafoclimáticas (ETr e Ks), quando aplicado em um solo contendo Imazapic;
- c) Avaliar o potencial do modelo neural ensemble, para explicar a correlação e influência de variáveis edafoclimáticas, quando aplicadas nos processos produtivos de um biocarvão.

2.3 ORGANIZAÇÃO DO TRABALHO

Este trabalho aborda na sua parte introdutória, a questão da disponibilidade alimentar mundial, sua importância para atender a uma crescente taxa demográfica, frente aos diversos desafios encontrados, como necessidade urgente de aumento da produtividade por hectare plantado, diversidade de culturas, qualidade nutricional, além dos impactos ambientais decorrentes. Retrata rapidamente, as soluções amparadas nas ferramentas tecnológicas, como aplicações de sistemas computacionais para a gestão fim a fim da produção (AP), a modelagem computacional para melhorias de processos ambientais, visando principalmente, investigação e aplicação de mecanismos de auxílio à atividade

agrícola, com menor impacto na degradação ambiental, como o biocarvão, amplamente pesquisado na literatura acadêmica.

A tese está organizada na forma de artigos. O primeiro capítulo realiza a modelagem computacional de variável ambiental influente nos processos agrícolas, a evapotranspiração real diária (ET_r). Modelagem utilizando algoritmos de aprendizagem de máquina investiga sua estimativa, calculada por variáveis secundárias. Este estudo é importante para viabilizar o uso da ET_r, quando da investigação de sua influência na produção de um biocarvão. Ele foi submetido como artigo a uma revista especializada internacional. No segundo capítulo, o comportamento e a influência de variáveis preditoras, utilizadas na estimação da condutividade hidráulica saturada (K_s), são avaliadas. Novamente, modelagem computacional com algoritmos de aprendizados de máquinas são testados, visando a otimização de um experimento realizado em campo, com valores das variáveis preditoras calculados em bancada. A determinação de valor estimado de K_s por métodos de IA viabilizou o estudo de sua influência na modelagem do processo produtivo de um biocarvão. Ele também foi submetido a uma revista especializada internacional. Já no terceiro capítulo, um modelo computacional neural artificial é desenvolvido, utilizando os dois primeiros algoritmos para estimar valores de ET_r e K_s. Eles alimentam um outro modelo neural (modelo ensemble), avaliando a influência destas variáveis na determinação de curvas isotermas. Estas isotermas determinam a capacidade adsortiva de um biocarvão, quando aplicado em solo contendo valores residuais de um herbicida. Este trabalho também foi submetido a um periódico internacional. O trabalho é encerrado com uma conclusão final.

Este trabalho procura contribuir para desenvolvimento de mais pesquisas com biocarvão, viabilizando aprofundamento da compreensão de seus mecanismos de formulação e atuação, inclusive para sua aplicação em sistemas aplicativos de agricultura de precisão, disponibilizando novas ferramentas para ampliação dos estudos da ciência ambiental.

3 REFERENCIAL TEÓRICO

3.1 BIOCARVÃO

A molécula do Imazapic (WANG et al., 2023) tem sido utilizada largamente em herbicidas para controle de plantas daninhas em culturas diversas. O principal efeito de sua aplicação ocorre sobre a enzima aceto lactato sintase, um importante elemento na biossíntese de aminoácidos por organismos vegetais, implicando na influência do crescimento vegetal. Seu ciclo de atuação ocorre no solo, sendo altamente solúvel em água, entre 30 e 90 dias, dependendo das variações ambientais. No Brasil, tem sido aplicada nos plantios de arroz irrigado (JUNKES et al., 2022), amendoim (DARAMOLA et al., 2024), e principalmente, em plantios de cana-de-açúcar (SARTORI et al., 2023), sendo ressaltado seu poder em ocorrências resistentes. No entanto, sua persistência em diversas concentrações no ambiente pode representar um risco ao ecossistema e à saúde humana, dada sua facilidade de transporte e fixação, além de sua contribuição nos processos de lixiviação. Ele pode incorrer em efeitos prolongados para sua biodegradação em solos ácidos, sob características de monocultivo e baixo teor de matéria orgânica. Sua aplicação repetitiva tem revelado ocorrência de biótipos altamente resistentes, e tem sido frequente o seu uso combinado a outros pesticidas (multissítios). Os efeitos decorrentes deste herbicida e de outros pesticidas, têm sido avaliados por técnicas de remediação, como a biorremediação microbiana, fitorremediação, dentre outros.

O biocarvão, composto rico em carbono, produzido por método de pirólise de resíduos orgânicos diversos, como descartes de colheita agrícola, industriais ou de esterco animal. Sua composição apresenta uma alta área de superfície de contato, estrutura porosa, além da presença de importantes e variados grupos funcionais. Estas características, dentre outras, o tornam ferramenta poderosa para aplicações diversas, sendo bastante pesquisado nas últimas décadas como imobilizador de poluentes, regulador de solos, ou sequestro de carbono. Em pesquisas de remediação pesticidas, o biocarvão tem sido avaliado como um adsorvente eficaz para grande variedade de contaminantes no solo, decorrentes da atividade agrícola, como fungicidas e herbicidas (Figura 1). O biocarvão também tem sido proposto como ferramenta eficaz na imobilização de Imazapic em solos, minimizando ocorrência de porções excedentes ou residuais, visando evitar aumento de contaminação de solos e de águas subterrâneas, impactando os serviços ambientais. O biocarvão é dependente da biomassa carbonácea, e das condições de seu processo produtivo, como temperatura de pirólise, taxa de aquecimento, tempo de residência. Dentre

os atributos-chave inerentes à sua formulação, que direcionam sua aplicação, destacam-se a área de superfície de resposta, o pH, a capacidade de troca catiônica e relações carbono/nitrogênio (C/N), composição química, teor de cinzas, dentre outros.

A maioria das pesquisas na literatura tem focado em adsorção de contaminantes, melhoria da fertilidade do solo e sequestro de carbono.

Figura 1 – Evolução das pesquisas com biocarvão



Fonte: Web of Science (2024)

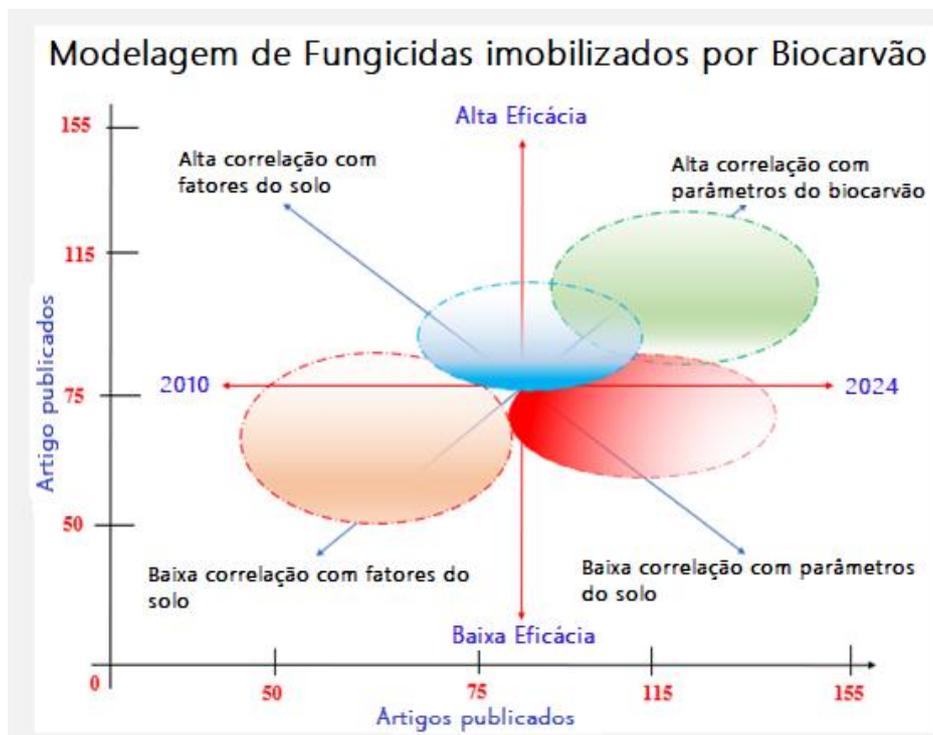
As pesquisas com adsorção de contaminantes, em particular, apresentam bons resultados, principalmente em métodos de remediação ou regulação de solos poluídos com agrotóxicos, como o Imazapic, frente a experimentos em bancada e aplicações de campo.

A eficácia do biocarvão na imobilização da molécula de Imazapic pode ser influenciada por uma série de fatores, incluindo técnicas de produção e o tipo de solo em que é aplicado. Portanto, estudar como os mecanismos e fatores que afetam a imobilização do Imazapic no solo, através da ação do biocarvão é importante, pois, além contribuir para compreensão da imobilização de moléculas poluentes no solo, facilita a análise e o desenvolvimento de novas estratégias que visem potencializar seu manejo em solos contaminados. Além disso, a cinética de libertação e o destino ambiental dos pesticidas imobilizados precisam ser cuidadosamente investigados para avaliar a sua estabilidade a longo prazo e os riscos potenciais. As interações entre o biocarvão e os componentes do solo, como matéria orgânica, minerais argilosos e microrganismos do solo, também devem ser consideradas para compreender a sua influência na imobilização de pesticidas.

Outro fator importante do potencial de uso do biocarvão diz respeito ao estudo das principais lacunas para sua ampla adoção como ferramenta de adsorção, como risco de contaminação residual, por uso de biomassa contendo metais pesados, efeitos de longo prazo no meio de aplicação, não plenamente fundamentadas, padronização de processo produtivo, falta de estudos prévios e adaptabilidade aos meios de aplicação, reduzindo sua efetividade. Outro importante fator a ser considerado, é seu custo produtivo e transporte, que dependendo de seu processo produtivo, pode variar de R\$ 5.000,00 a R\$ 20.000,00 por tonelada aplicada.

Pesquisas em desenvolvimento na literatura apontam para que atenuar a maioria destes problemas, algumas medidas devem ser adotadas, tais como: investigar a formulação de biocarvões funcionais e formulações combinadas, inclusive fertilizantes; aplicação integrada com sistemas de agricultura regenerativa; e estímulo e ampliação de políticas voltadas para sua disseminação e aquisição. Outra importante linha de pesquisas é a utilização de técnicas de inteligência artificial (Figura 2), para investigar o biocarvão em todas as suas nuances, principalmente com técnicas envolvendo a aplicação de modelagem computacional, apoiadas por algoritmos de aprendizado de máquina (GUPTA et al., 2023).

Figura 2 – Direção das pesquisas de biocarvão com modelagem de Inteligência Artificial



Fonte: Web of Science (2024)

Estimativas de organismos europeus estimam que o mercado de biocarvão ultrapassará os U\$S 3 bilhões até 2030, estimulado por políticas de sistemas agrícolas sustentáveis.

3.2 MODELAGEM DE PROCESSOS AMBIENTAIS

A modelagem com o uso de Inteligência Artificial (IA) tem trazido grandes avanços para a pesquisa, dada sua capacidade em manipular sistemas complexos, com grandes quantidades de dados, de alta variabilidade espacial, e com variáveis não lineares, mas com correlação, como são os processos ambientais. Os modelos aplicados de IA possuem, ainda, robustez e precisão, pois conseguem incorporar métodos estatísticos e computacionais para disponibilizar recursos de análise explicativas e avaliação de qualidade na predição e otimização de fenômenos ambientais diversos, como avaliação de águas, poluição do solo, dispersão de contaminantes, e aqueles com mudanças climáticas. É possível, ainda, integrar diferentes bases e estruturas de dados. Modelagem com IA tem superado a modelagem tradicional estatística, por capturar padrões ocultos nos dados, identificação de padrões e delineamento de forma online. Na literatura acadêmica têm sido apresentados estudos diversos, com múltiplos objetivos, utilizando algoritmos de aprendizado de máquina de acordo com as características das variáveis envolvidas. Dentre os mais utilizados em processos ambientais, destacam-se as Redes Neurais Artificiais - RNA (DIYKH et al., 2024), Random Forest – RF (KANG et al., 2023), Support Vector Machine – SVM (FAOUZI et al., 2023), e XGBoost (TARWIDI et al., 2023).

Machine Learning (ML) e Deep Learning (DL) são algoritmos de aprendizado de máquina, como redes neurais artificiais e redes neurais profundas, que, associadas ou comparativamente a árvores de decisão (RF) e máquinas de vetores de suporte (SVM), são frequentemente usados na modelagem ambiental. Estes algoritmos podem identificar padrões complexos em grandes conjuntos de dados, essenciais para compreender processos ambientais.

As Redes Neurais Convolucionais (CNNs) são métodos computacionais com inteligência artificial especialmente úteis na análise de imagens de satélite e dados espaciais, tendo sido aplicadas para rastrear mudanças ambientais, como desmatamento e expansão urbana.

O processamento de Linguagem Natural (PLN) evoluiu intensamente na última década, hoje sendo utilizada para analisar dados textuais ambientais, como relatórios de impacto ambiental e publicações científicas, permitindo uma melhor compreensão das tendências e preocupações ambientais.

Dentre os principais desafios e limitações está a obtenção de dados de qualidade. A falta de dados rotulados e representativos é um desafio significativo, pois os modelos de IA dependem fortemente da qualidade e quantidade dos dados para sua precisão e confiabilidade. Outro importante fator na modelagem com IA é a interpretabilidade dos modelos de dados com IA, que melhor representem a variabilidade espacial dos dados, especialmente tratados com redes neurais profundas, frequentemente vistos como "caixas-pretas", dificultando a compreensão de como as decisões são tomadas. A integração do conhecimento específico do domínio de dados explorados é imperativa com o conhecimento específico do domínio ambiental, visando garantir que os modelos sejam realistas e aplicáveis.

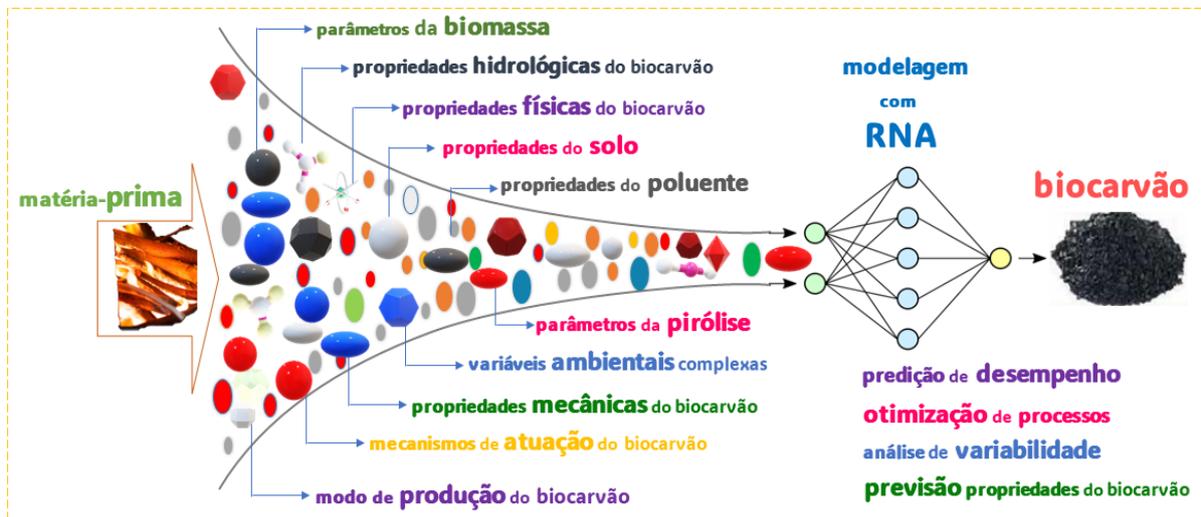
Dentre as aplicações práticas de IA estão as que tratam sobre monitoramento ambiental, onde a IA é usada para monitorar a qualidade do ar, água e outros parâmetros ambientais, fornecendo dados em tempo real para tomada de decisão. Da mesma forma, a previsão de eventos climáticos vem sendo tratada com maior frequência com modelos preditivos em grandes volumes de dados históricos, com fim de prever eventos extremos, como inundações e secas, permitindo ações preventivas. Da mesma forma, diversos estudos vêm sendo realizados com utilização de IA na conservação da biodiversidade, na identificação e monitoramento de espécies em risco, utilizando imagens de satélite e dados de sensores. A integração de multi-modelos, combinação de diversas plataformas matemáticas e estatísticas, em um único ambiente de modelagem de processos com suporte em IA, pode fornecer insights mais abrangentes e precisos sobre os processos ambientais.

A aplicação de IA na modelagem de processos ambientais oferece possibilidades notáveis para melhorar a compreensão e gestão do meio ambiente. No entanto, é fundamental abordar os desafios relacionados com a qualidade dos dados, interpretabilidade dos modelos e integração de conhecimento específico do domínio. O avanço promovido nesta área promete não apenas melhorar as ferramentas de modelagem ambiental, mas também contribuir para uma gestão mais sustentável e eficaz dos recursos naturais.

O aprimoramento da modelagem de processos envolvendo biocarvão é percebido por novas pesquisas, que agora associam aplicação das redes neurais artificiais (RNA) a

outros modelos de IA, visando aprofundar conhecimento nos padrões dos dados e sua interpretabilidade (Figura 3), melhorando entendimento de seu processo de formulação e atuação.

Figura 3 – Parâmetros do biocarvão modelados por Aprendizado de Máquina (IA).



Fonte: Autor (2024)

A utilização de modelos de IA na modelagem de biocarvões representa uma abordagem inovadora e eficiente, capaz de abordar tanto aspectos técnicos quanto ambientais e econômicos do processo. Essa abordagem multidimensional é crucial para desenvolver práticas sustentáveis e eficazes de produção de biocarvão, contribuindo para a mitigação das mudanças climáticas e promovendo uma agricultura sustentável. A pesquisa contínua e o desenvolvimento de modelos mais precisos e robustos são essenciais para avançar neste campo.

4 CAPÍTULO 1 – PREDIÇÃO DE ÍNDICE DE EVAPOTRANSPIRAÇÃO REAL DIÁRIA AVALIADA POR MODELOS DE INTELIGÊNCIA ARTIFICIAL

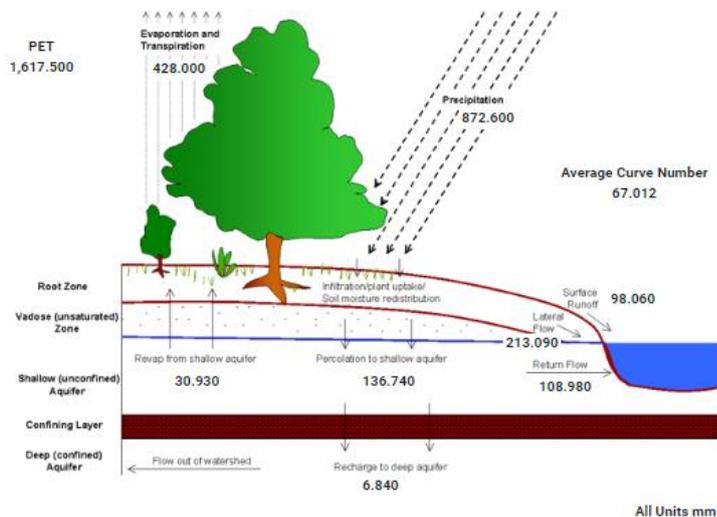
PONTOS IMPORTANTES

Ampliar conhecimento sobre evapotranspiração, diferenças metodológicas nos métodos computacionais de inteligência artificial utilizados para suas estimativas;

Avaliar acurácia e precisão dos sistemas de estimativas, visando otimização de calibração e disponibilização de informações;

Destacar aplicação de inteligência artificial como opção atrativa na modelagem de processos ambientais, em procedimentos de otimização de dados envolvendo variáveis espaciais, ou avaliação qualitativa de informações disponibilizadas;

RESUMO GRÁFICO



4.1 INTRODUÇÃO

O fenômeno da evapotranspiração é definido como um processo que promove a transferência da água da superfície, oriunda de ocorrências de evaporação e da transpiração das plantas para a atmosfera (DOUNA et al., 2021; HAO et al., 2022). A evapotranspiração real diária (ET_r) é uma variável crucial em sistemas hidrológicos, atuando como um vetor de transporte para devolver mais de 60% das precipitações para a atmosfera (DAYAL et al., 2021; SU et al., 2022; BAFTI et al., 2024), sendo crítica para a compreensão e mapeamento dos ciclos de água em atividades diversas, como agrícolas e sistemas de irrigação.

A medição e avaliação da ET_r têm sido pesquisadas através de técnicas direta e indireta, além da aplicação de modelos computacionais. Dentre os métodos experimentais

diretos, destacam-se as medições realizadas *in situ*, através da utilização de equipamentos como lisímetros de pesagem, cintilômetros de larga abertura, ou sistemas de covariância de vórtices, dentre outros, com alto custo de aquisições dos dados experimentais (SENAY et al., 2023; BAFTI et al., 2024; CHAKHAR et al., 2024), exigindo tempo e pessoal qualificado, implicando em escalas locais, podendo incorrer em imprecisões de medições.

Entre os métodos indiretos destacam-se os baseados em balanço de energia, como Bowen Ratio e Scintilômetro (BASAKIN et al., 2023; DE SOUZA et al., 2023), envolvendo medições por outras variáveis como radiação líquida, fluxo de calor no solo, fluxos de calor latente, requerendo, algumas vezes, coleta física *in situ*, além de exigirem precisão, também com custos adicionais. As técnicas de medição envolvendo modelos empíricos e físicos aplicam principalmente as equações de Penman-Monteith, Priestley-Taylor e Modelos de Superfície Terrestre – LSMs (DA SILVA et al., 2022; DE-VILLE et al., 2024), com muitos estudos publicados na literatura. Outro método indireto com grande acervo em pesquisas, envolve as técnicas de sensoriamento remoto (SR), com produtos gerados por várias plataformas satelitais, utilizando modelagens com algoritmos como o SEBAL, Metric, e Índices de Vegetação, como NDVI (ALSENJAR et al., 2023; DANDA et al., 2023; GONG et al., 2023; MA et al., 2023; BURCHARD-LEVINE et al., 2024; BUTLER et al., 2024). Por fim, os métodos com modelagem computacional, baseados em técnicas aplicadas com algoritmos de inteligência artificial (IA), vêm sendo bastante investigados na literatura, principalmente, com predição ou regressão de resultados por aprendizado de máquina. Apresentam boa capacidade de aproximação, aceitabilidade de ampla gama de tipos de dados de entrada, manipulando múltiplas variáveis, com êxito em capturar padrões temporais e espaciais de difícil replicabilidade, reduzindo a complexidade e incrementando a interpretabilidade de modelagem (DOUNA et al., 2021; GONZÁLEZ et al., 2023; KIM et al., 2024; MOZAFFARI et al., 2024).

Recentes pesquisas na literatura têm sido apresentadas com redes neurais profundas (FANG et al., 2022; BAFTI et al., 2024; DIYKH et al., 2024; KIM et al., 2024), de forma comparativa ou mesmo combinadas a algoritmos poderosos de aprendizado (HAO et al., 2022; MAMUN et al., 2024), visando capturar interações complexas de diversos fatores, que impliquem na medição precisa e confiável de ETr, contribuindo com flexibilidade e aprimoramento de resultados de estimação. A maioria dos modelos de previsão projetados, tem utilizado séries temporais de dados, oriundas de imagens satelitais (BEYER et al.,

2023; GUO et al., 2024), ou mais recentemente, de drones (MOKHTARI et al., 2023; TURMAN et al., 2024).

Previsões de ETr com séries temporais podem incorrer em imprecisões causadas eventualmente por ruídos diversos, erros atmosféricos, baixa resolução, lacunas temporais no imageamento satelital ou por drones. Algumas variáveis, como NDVI, por não serem medidas diretamente pelos sensores satelitais, mas sim estimadas indiretamente, com modelos como o SEBAL, Metric, EEFlux ou Penman-Monteith modificado, utilizando-se outras variáveis de entrada, como temperatura da superfície, albedo, NDVI e radiação solar, podem ter sua estimativa também influenciada por imprecisões destes métodos. Visando atenuar estes problemas, pesquisas com modelos de redes neurais profundas (*deep learning*) têm sido realizadas, com boas aproximações, capturando variações temporais e espaciais, utilizando as informações dos canais das imagens, ou mesmo técnicas associadas, como transformações de curvelets (DIYKH et al., 2024; KIM et al., 2024).

A combinação e diversidade de estratégias de modelagens de processos ambientais permite ampliação ou mesmo customização das pesquisas direcionadas, tornando interessante, por exemplo, a aplicação de aprendizado profundo (DL), associando variáveis de entrada ao objetivo específico de interesse da predição de ETr, suportado por um modelo de IA robusto, que atenua as imprecisões associadas, customizando recursos para pesquisa. Ao estimar o espaço amostral de ETr, tendo como variáveis de entrada ETr e NDVI, geradas do produto MOD16A2, com resolução temporal de 8 dias e espacial de 500 m, embora não garanta análises adequadas para estudos regionais ou de monitoramento sazonal, é viável para gestão de recursos hídricos, ou monitoramento de bacias hidrográficas, em estudos de larga escala, identificando áreas sob estresse hídrico. Em aplicações de cultivos agrícolas, a correlação NDVI-ETr para um ETr estimado pode subsidiar sistemas inteligentes de irrigação, sendo contraindicada em regiões áridas, dependendo de calibração local ou integração com dados complementares. Outra destinação possível para predições de ETr, derivadas de correlação NDVI-ETr, pode ser a investigação de sua influência como insumo crítico na formulação de bioprodutos, que mantenham relações com a biomassa, qualidade do solo, ou mesmo otimização de produtos, como na produção de biocarvões, produzidos especificamente para dada região (clima, solo, ciclo de águas), associando pH, carbono orgânico no solo, ou outras propriedades qualitativas essenciais.

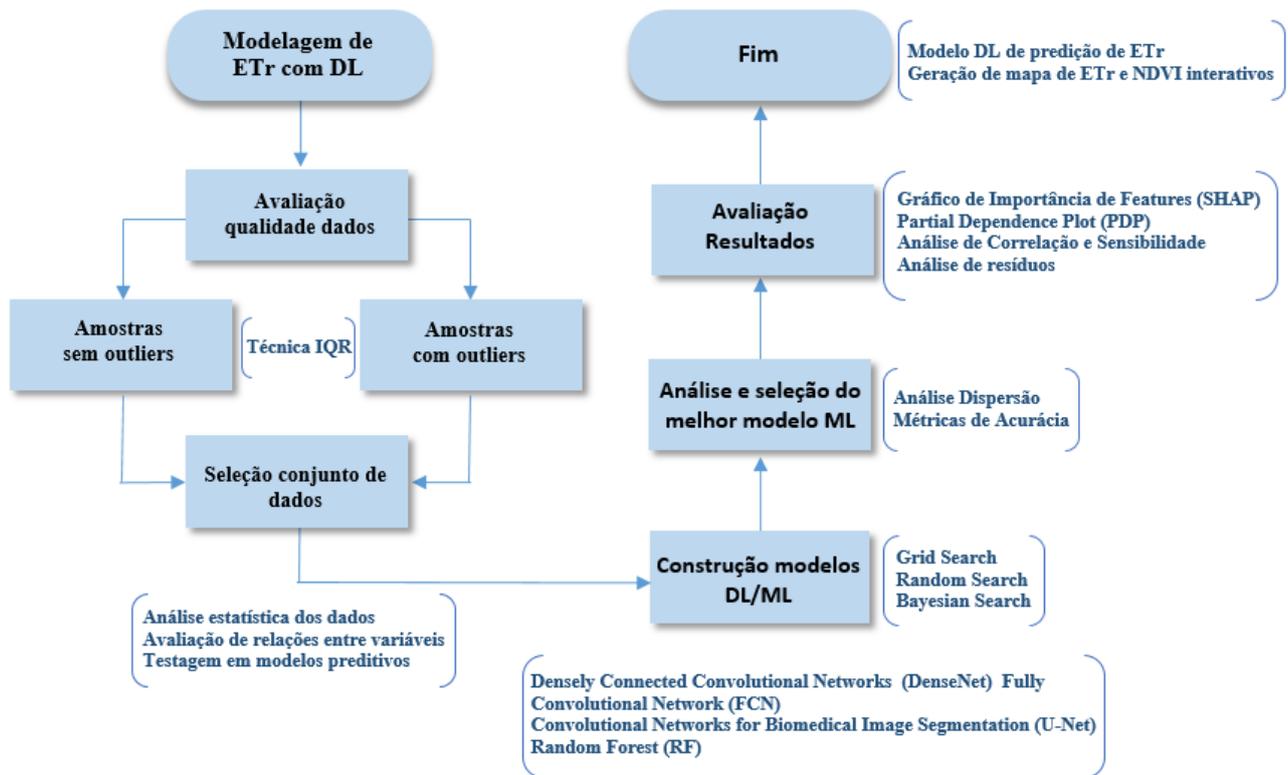
Assim, o principal objetivo desta pesquisa foi avaliar a qualidade da predição da evapotranspiração real diária, a partir de um pequeno conjunto de dados de NDVI e ETr, obtidos de informações satelitais, avaliando a robustez de quatro modelos de aprendizado de máquina profundo, desenvolvidos para capturar as relações complexas e variabilidade espacial das amostras, de maneira eficiente e com melhor grau de confiabilidade, flexibilizando áreas de estudos, sistematizando viabilidade dos dados estimados, permitindo sua utilização em outras pesquisas, ou alimentação de aplicações de monitoramento e gerenciamento de recursos hídricos.

O artigo está estruturado em quatro etapas. A segunda etapa, de Materiais e Métodos, detalha a sequência de desenvolvimento da pesquisa, os procedimentos de geração de dados e descrição da região de estudo, construções dos modelos de IA, definição de melhores hiperparâmetros para os modelos e o desempenho dos modelos avaliados. A etapa de Resultados e Discussões compreende breve análise do conjunto de dados originais, análises dos resultados da predição pelo melhor modelo de IA e análises explicativas dos resultados obtidos. A etapa de Conclusões resume o estudo e oferece recomendações para estudos futuros.

4.2 MATERIAIS E MÉTODOS

Os dados geoespaciais foram obtidos visando avaliar quatro modelos de inteligência artificial, descritos a seguir, para estimar evapotranspiração real diária (ET_r). Duas variáveis foram mensuradas para esta estimação, sendo NDVI e ET_r como entrada para os modelos, e uma variável-resposta (ET_r estimada). Um conjunto de mil e cinquenta e cinco amostras foi avaliado por métodos estatísticos, identificando a existência de pontos extremos (*outliers*) para cada variável de entrada. O impacto de manutenção ou poda destes pontos extremos foi avaliado. Na sequência, quatro modelos de inteligência artificial, sendo três de aprendizado de máquina profundo (DL) e um de aprendizado (ML), comumente descritos na literatura acadêmica para modelagem de processos ambientais, inclusive determinação de NDVI e ET_r, além da natureza dos dados, foram desenvolvidos e avaliados com este conjunto de dados. Visando garantir a seleção do melhor modelo de IA para a predição de resultados de modelagem, uma padronização de procedimentos para a construção das arquiteturas foi aplicada, respeitando-se suas especificidades, consistindo em parametrização automática de seus atributos-chave (definição de hiperparâmetros) por métodos estatísticos de busca, seguido de avaliação de suas métricas de resultados. O modelo com melhores métricas de acurácia foi utilizado para predição da modelagem da evapotranspiração diária do solo de estudo, além de gerar uma proposta de mapa interativo de maior representatividade espacial das variáveis preditoras do modelo (Figura 1).

Figura 1 – Descrição das etapas aplicadas para avaliação dos modelos de DL e ML para a estimação da evapotranspiração real diária (ET_r).



Fonte: Autor (2024).

4.2.1 Geração de dados e delineamento experimental

Os dados experimentais de NDVI e ET_r utilizados nesta pesquisa foram gerados pelo sistema MOD16A2, da NASA (MOD16A2, 2024). Ele fornece estimativas de evapotranspiração derivadas do modelo de balanço da vegetação da área de estudo, com resolução espacial de 500 metros e intervalo temporal de 8 dias. O período escolhido compreendeu o ano de 2023, garantindo uma amostragem mínima das variações ao longo das estações do ano.

A região de estudo compreende uma área de origem agrícola, com aproximadamente 2.500 m², comumente utilizada para pastagem de gado, onde a vegetação cultivada é a *Brachiaria decumbens* Stapf. A área fica localizada na fazenda Riacho do Papagaio, município de São João (Figura 2), na mesorregião do Agreste Meridional, no estado de Pernambuco, de coordenadas geográficas 8°52'30" S, e 36°22'00" W, em altitude de 705 m. O clima observado é predominantemente o As', sendo equivalente a clima quente e

úmido, conforme classificação de Köeppen (BECK et al., 2018). O período chuvoso é previsto para o período de outono/inverno, com início em dezembro/janeiro e término até setembro. A precipitação média anual registrada na literatura para a região é de aproximadamente 1.310 mm.

As imagens de satélite foram processadas na plataforma Google Earth Engine (GEE, 2024), que permite análise em nuvem de grandes volumes de dados de sensoriamento remoto, com armazenamento otimizado e ampla curadoria de coleções. As variáveis de interesse (entradas dos modelos de IA) incluíram índices espectrais e estimativas de evapotranspiração obtidas por esquema K_c - E_{To} ajustado por estresse hídrico. Adotou-se a forma simplificada

$$ET_r = E_0 \times K_c \times K_s$$

Onde:

- E_{To} é a evapotranspiração de referência (Penman–Monteith FAO-56, mm dia^{-1}),
- K_c é o coeficiente de cultura,
- K_s é o coeficiente de estresse hídrico ($0 \leq K_s \leq 1$).

Nessa formulação, ET_r fornece uma estimativa conservadora da evapotranspiração de cultura sob déficit hídrico, pois ($0 \leq K_s \leq 1$). Nessa formulação, ET_r fornece uma estimativa conservadora da evapotranspiração de cultura sob déficit hídrico, pois $K_s \leq 1$ reduz o termo $K_c E_{To}$ quando a água no solo é limitante (ALLEN et al., 1998). Aqui foi utilizada a versão agregada (um único K_c) por concisão e consistência com a etapa de IA.

O NDVI foi calculado em GEE a partir das reflectâncias de infravermelho-próximo (NIR) e vermelho (RED):

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$$

Onde:

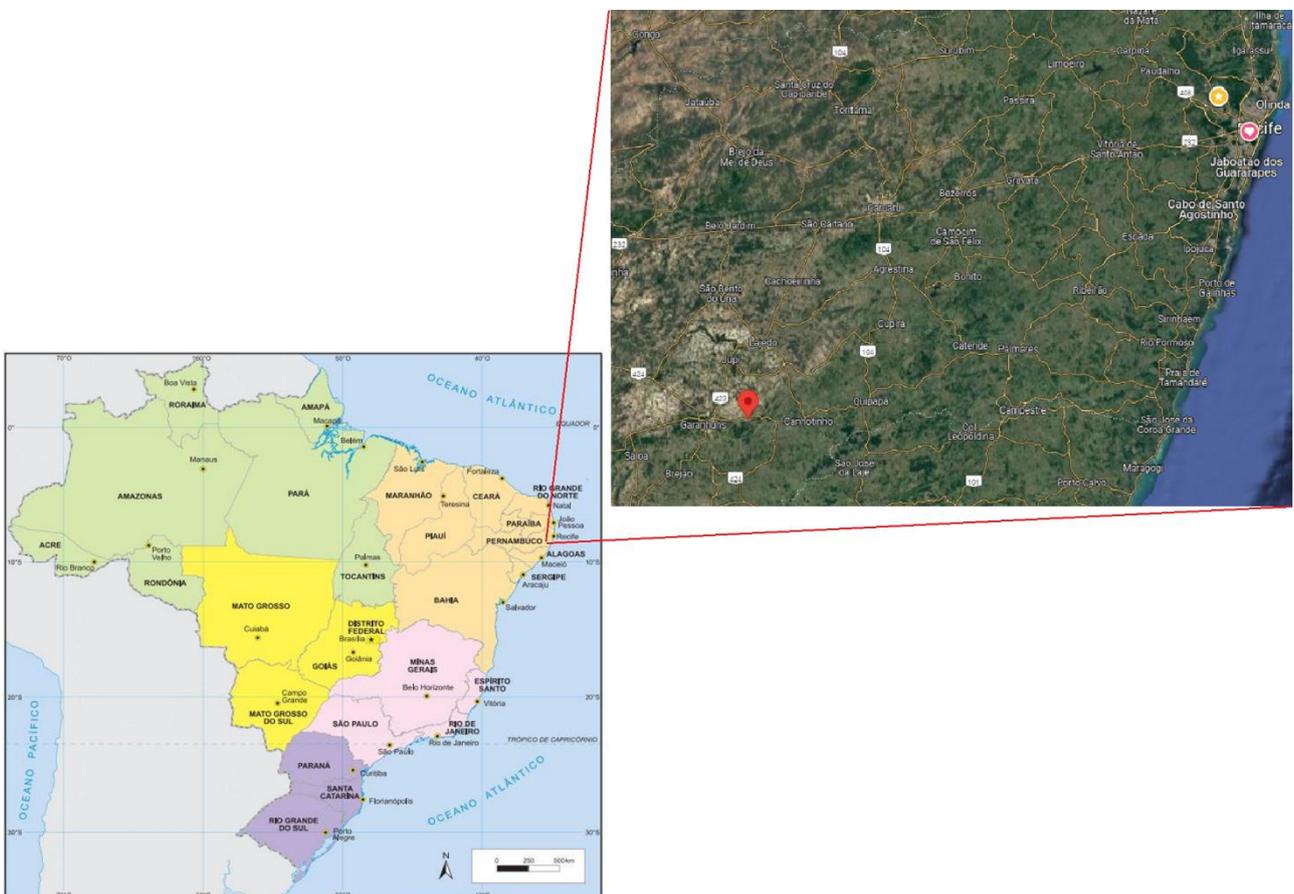
- NIR representa a reflectância na faixa do infravermelho próximo;
- RED representa a reflectância na faixa do vermelho;

e varia de -1 a $+1$, servindo como proxy de vigor e densidade foliar (ROUSE et al., 1974). O NDVI não estima ET diretamente; ele é utilizado como variável explicativa para calibrar K_c ou por *scaling* entre $K_{c,\min}$ e $K_{c,\max}$ e/ou para derivar fração de cobertura e LAI usados em

modelos de balanço de energia (BAUSCH & NEALE, 1987; ALLEN et al., 2007; BASTIAANSEN et al., 1998).

Após o processamento no GEE, NDVI e ETr foram amostrados e recortados à área de estudo (máscara vetorial), e exportados para QGIS® (QGIS, 2024) para conferência cartográfica. A tabela final para análise estatística consistiu em duas planilhas (n = 1.055 pontos cada): uma contendo os pontos e valores de ETr e outra contendo os pontos e valores de NDVI. A estrutura temporal/espacial seguiu o delineamento experimental definido para os modelos de IA.

Figura 2 – Localização geográfica da área de estudo, São João/PE.

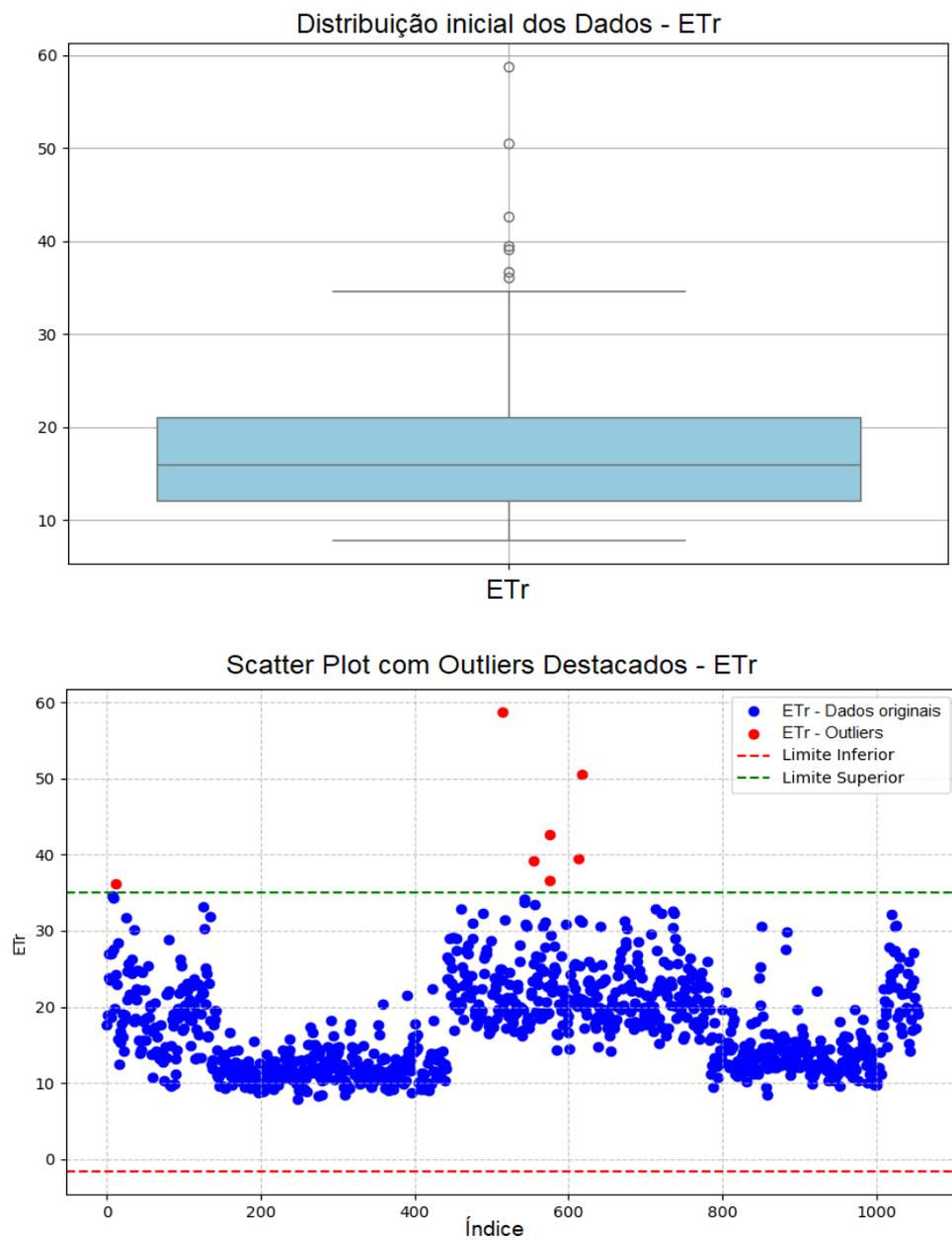


Fonte: Autor (2024)

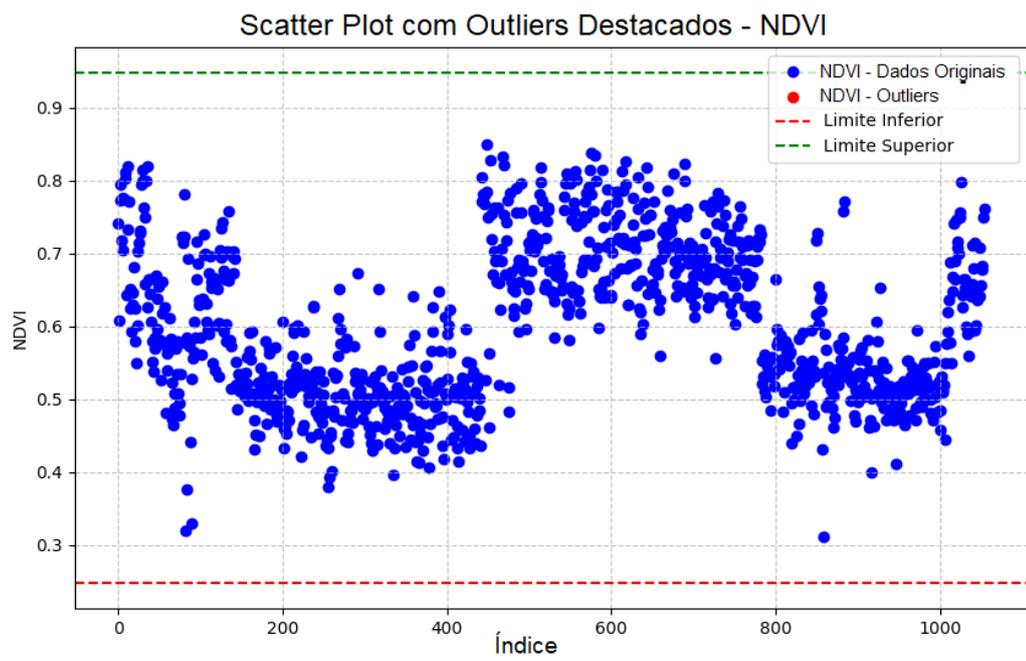
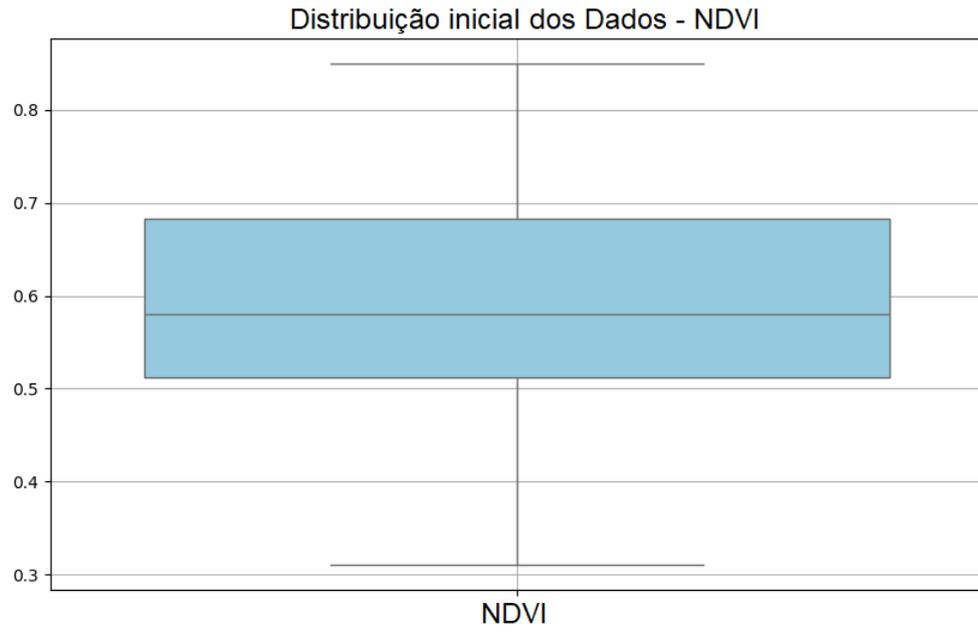
4.2.2 Tratamento de dados experimentais

O conjunto de dados geoprocessados contendo duas colunas (ETr e NDVI) de 1.055 amostras cada, foi submetido à análise de *outliers*, utilizando-se a técnica de Interquartil - IQR (KORKMAZ, 2023), permitindo visualizar 7 pontos extremos (Figura 3a), sendo ETr = 7 pontos extremos, e NDVI sem nenhuma ocorrência (Figura 3b).

Figura 3 – Box plot dos pontos extremos por variável, no conjunto amostral original.



(a)



(b)

Fonte: Autor (2024)

O impacto da poda de *outliers* na distribuição e representatividade das relações entre variáveis foi procedido através de análise exploratória de dados (Tabela 1).

Tabela 1 – Análise descritiva do conjunto amostral original e estimado

	Média	Mínimo	Máximo	Intervalo	Assimetria	Curtose
NDVI original	0,4442	0	1	1	0,2091	-1,655
NDVI estimado	0,4416	0,0685	0,9487	0,8802	0,2395	-1,6982
ETr original	0,4492	0	1	1	0,2765	-1,6406
ETr estimado	0,5043	0,1	1	0,9	0,2765	-1,6406

Fonte: Autor (2024)

4.2.3 Construção dos modelos de aprendizado de máquina

Quatro modelos de inteligência artificial foram selecionados, tendo como premissa, pesquisas na literatura sobre modelagem de processos ambientais utilizando algoritmos de aprendizado de máquina, além das características dos dados experimentais (MUÑOZ-CARPENA et al., 2023). Os algoritmos de rede neural profunda DenseNet (DIYKH et al., 2024), FCN (FANG et al., 2022) e U-Net (BAFTI et al., 2024), além do algoritmo de aprendizado Random Forest (KANG et al., 2023; REZAEI et al., 2023; MAMUN et al., 2024) foram desenvolvidos para a modelagem de previsão ETr, buscando capturar as relações entre as variáveis de entrada e saída, utilizando *frameworks* apropriados para cada algoritmo de IA na linguagem de programação Python®. Visando facilitar a robustez de cada modelo construído, assim como possibilidade de comparação de resultados, todos utilizaram estratégias similares, com etapas claras de: leitura de dados: (i) pré-processamento: normalização com função *StandardScaler*, divisão de dados com 20% dos dados para teste, transformação Yeo-Johnson); (ii) treinamento: otimização de hiperparâmetros (seleção de métodos de buscas Grid Search, Random Search e Bayesian Search), regularização L1/L2, parada antecipada, atribuição de tamanho camadas ocultas, taxa de aprendizado, avaliação (métricas MAE, MSE, RMSE, R² e Coeficiente de Correlação); avaliação de robustez (validação cruzada com 5 divisões); (iii) avaliação de resultados (análise de resíduos com histogramas e gráficos de dispersão, explicabilidade com PDP e SHAP); (iv) avaliação de resultados.

4.2.4 Hiperparâmetros dos modelos de ML por mecanismos de busca

Três métodos de busca usualmente utilizados na literatura, Grid Search (TARWIDI et al., 2023), Random Search (SHAN et al., 2023), e Bayesian Search (ROBERT; ROUSSEAU, 2023) foram aplicados em todas as arquiteturas desenvolvidas, visando definição automática do melhor conjunto otimizado de hiperparâmetros. Cada modelo de IA foi elaborado para utilizar hiperparâmetros apropriados às suas características funcionais (VON KROGH et al., 2023). Dentre os principais atributos avaliados em cada modelo de IA, destacam-se: Redes Neurais Artificiais (*hidden_layer_sizes*, *activation*, *alpha*, dentre outros), Random Forest (*n_estimators*, *max_depth*, *min_samples_split*, e outros).

4.2.5 Desempenho dos modelos de Aprendizado de Máquina

Para classificação e seleção do melhor modelo de IA na estimação de ETr, foram calculadas as métricas de MSE, MAE, RMSE, R² e CC de cada algoritmo de aprendizado de máquina, visando avaliar comparativamente performance.

4.3 RESULTADOS E DISCUSSÃO

A preocupação com perda de dados e informações relevantes correlacionadas é uma realidade importante nos processos de modelagem de dados com inteligência artificial, principalmente em fenômenos complexos, envolvendo variáveis com variabilidade espacial e temporal, como os ambientais. Ela é quantificada comumente como uma porcentagem da variância dinâmica do sistema (MUÑOZ-CARPENA et al., 2023; SHREIM et al., 2024). A lacuna de dados pode comprometer significativamente a qualidade da modelagem, além da capacidade preditiva dos modelos de IA aplicados, influenciando a captura de padrões, tendências e relações implícitas entre as variáveis preditoras e alvo, prejudicando a capacidade de generalização. Assim, estratégias de imputação de dados, validação cruzada rigorosa, seleção de variáveis e uso de métodos de explicabilidade devem ser implementadas, garantindo um equilíbrio entre acurácia e interpretabilidade.

4.3.1 Análise do conjunto de dados originais

A distribuição da ETr no conjunto original apresenta assimetria à direita, com maioria dos valores concentrada entre 1,74 e 2,91 mm/dia (faixa interquartilica, Q1–Q3) e 80%–90% da massa dentro de 1,56 a 3,42 mm/dia (p10–p90). Observam-se pontos extremos acima do limite superior do boxplot $\approx 4,65$ mm/dia, alcançando máximos entre 5,19 e 7,68 mm/dia. Esses extremos podem refletir regimes hidrometeorológicos de alta demanda evaporativa (combinações de radiação/temperatura/vento com baixa umidade), heterogeneidades espaço-temporais no balanço de energia e/ou incertezas de processamento. O padrão ao longo do índice (Fig. 3a, parte inferior) sugere segmentos temporais com regimes distintos de atmosfera e umidade do solo, compatíveis com alta variabilidade; como os outliers não se concentram em um único trecho, são plausivelmente parte do suporte físico do processo, e não mero ruído aleatório..

Para o NDVI, a distribuição é mais homogênea (Fig. 3b), com valores predominantemente entre 0,3 e 0,8, ausência de caudas longas e escassa ocorrência de extremos; tal comportamento é coerente com um dossel vegetal com cobertura e vigor relativamente estáveis ao longo do período, servindo como preditor robusto para a modelagem de ETr.

A análise descritiva (Tabela 1) quantifica essas observações. Para o NDVI, a média estimada pela rede (0,4416) permaneceu muito próxima da média original (0,4442), indicando baixo viés. O mínimo estimado foi maior (0,0685 vs 0,0), enquanto o máximo estimado foi ligeiramente menor (0,9487 vs 1,0) e o intervalo reduziu (0,8802 vs 1,0). Em conjunto, esses números mostram que o modelo evitou valores extremos nas duas pontas e produziu uma distribuição mais compacta, sem alterar substancialmente as estatísticas centrais: a assimetria aumentou apenas de 0,2091 para 0,2395 e a curtose manteve-se negativa (–1,655 para –1,6982), evidenciando platicurtose e compatibilidade com as inspeções visuais (QQ/box).

Para a ETr, a média estimada (0,5043) foi superior à média original (0,4492), diferença absoluta de 0,0551, caracterizando leve viés positivo. O mínimo estimado subiu de 0,0 para 0,1, o máximo permaneceu em 1,0, e o intervalo reduziu de 1,0 para 0,9. Assim como no NDVI, há compressão de amplitude nas caudas inferiores, ao passo que o extremo superior foi preservado; a assimetria (0,2765) e a curtose (–1,6406) permaneceram idênticas às do conjunto original reportado na Tabela 1, indicando que a forma global da distribuição (assimetria moderada e platicurtose) foi mantida, apesar do encurtamento do suporte

inferior. Esse comportamento é típico de modelos com regularização e perdas que penalizam extrapolações e reduzem a variância sem sacrificar o erro médio.

Do ponto de vista preditivo, os testes conduzidos com o modelo escolhido (DenseNet), com e sem poda de outliers, não evidenciaram degradação relevante das métricas (Tabela 2), sugerindo robustez da arquitetura ao tratamento de caudas. A estabilidade das estatísticas de forma (assimetria/curtose) e a redução moderada da amplitude ajudam a explicar por que as métricas globais (RMSE/MAE/R²/CC) se mantêm altas mesmo na presença de valores extremos: a rede captura padrões dominantes e regimes recorrentes e, ao mesmo tempo, amortece excepcionalidades na parte baixa da distribuição. Ainda assim, dados os indícios de regimes raros no topo (35–60 mm/dia), recomenda-se não podar sistematicamente os extremos; em vez disso, convém estratificar por faixas (baixa/média/alta demanda) nas avaliações e, quando pertinente, aplicar ajustes dirigidos (p. ex., calibração por quantis na cauda inferior) para preservar suporte físico sem induzir overfitting.

Em síntese, o conjunto original de ETr mostra alta variabilidade com cauda superior relevante, ao passo que o NDVI é estável e informativo. A versão estimada pelo modelo mantém as estatísticas centrais e a morfologia da distribuição, com leve aumento na média da ETr e menor amplitude, efeitos compatíveis com regularização e controle de extrapolações. Como o objetivo era verificar se os padrões gerais de evapotranspiração e a dinâmica da vegetação seriam reconhecidos e preservados, os resultados confirmam consistência estatística e coerência física, sem dependência crítica da poda de outliers.

Ressaltando que o objetivo geral da modelagem era verificar se os padrões gerais de evapotranspiração e vegetação foram capturados, o que foi confirmado.

4.3.2 Performance dos modelos de IA

O método de pesquisa de atributos *Bayesian Search* foi o de melhor desempenho geral, comparativamente a *Random Search* e *Grid Search*, nos quatro modelos testados, com os menores valores de MSE e RMSE e maiores valores de R² e CC.

Os resultados obtidos pelos modelos de IA foram próximos, do ponto de vista de valores de desempenho (Tabela 2). Análises estatísticas realizadas por meio do teste ANOVA (Kruskal-Wallis) não indicaram diferença significativa entre os métodos, sugerindo que as

abordagens obtiveram resultados comparáveis. O melhor desempenho foi alcançado pelo modelo DenseNet, seguido de FCN, RF e U-Net.

Tabela 2 – Classificação por desempenho dos modelos de IA treinados com conjunto de dados sintéticos.

		MSE	RMSE	MAE	R ²	CC
DenseNet	Bayesian Search	0,002	0,046	0,038	0,98	0,99
FCN	Bayesian Search	0,008	0,089	0,075	0,95	0,97
RF	Bayesian Search	0,035	0,187	0,15	0,85	0,92
U-Net	Bayesian Search	0,097	0,311	0,26	0,7	0,84

Fonte: Autor (2024)

A análise dos resultados de acurácia dos modelos de IA (Tabela 2) indicou a rede neural profunda DenseNet como melhor desempenho geral, evidenciado pelos menores valores de acurácia (mse, rmse, mae), e maiores coeficientes de correlação (CC e R²). O método de otimização de hiperparâmetros que apresentou melhor resultado em todos os modelos de IA foi o Bayesian Search, garantindo maior eficiência na busca pelos atributos ideais para cada modelo.

A rede DenseNet possui estrutura altamente conectada, realizando passagem de forma eficiente das informações entre suas camadas, facilitando a extração de padrões espaciais e temporais nos dados modelados, podendo explicar o fato de apresentar melhores resultados de acurácia. A arquitetura aplicada também foi favorecida, dado sua característica de densidade de conexões e uso adequado de regularização, minimizando ocorrência de *overfitting*, melhorando sua capacidade preditiva ante novos dados.

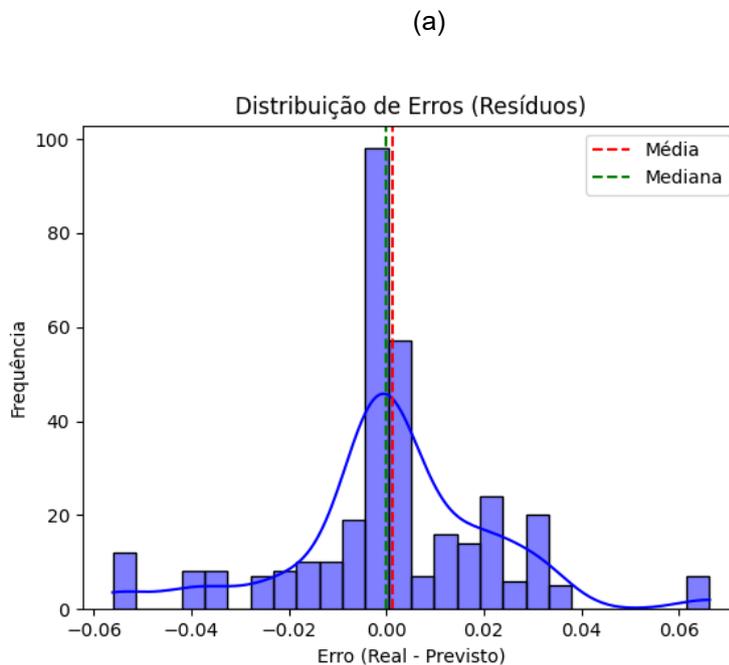
A rede FCN obteve bom desempenho, mas apresentou erro um pouco maior, dada sua característica de sua arquitetura, voltada para aprendizado espacial, mesmo tendo ótimo potencial para capturar características locais, o que pode não ter sido suficiente para generalizar a relação entre NDVI e ETr. Apesar disso, como apresentou bons resultados de R² (0,95) e CC (0,97), segue sendo alternativa válida para este tipo de baixa complexidade computacional.

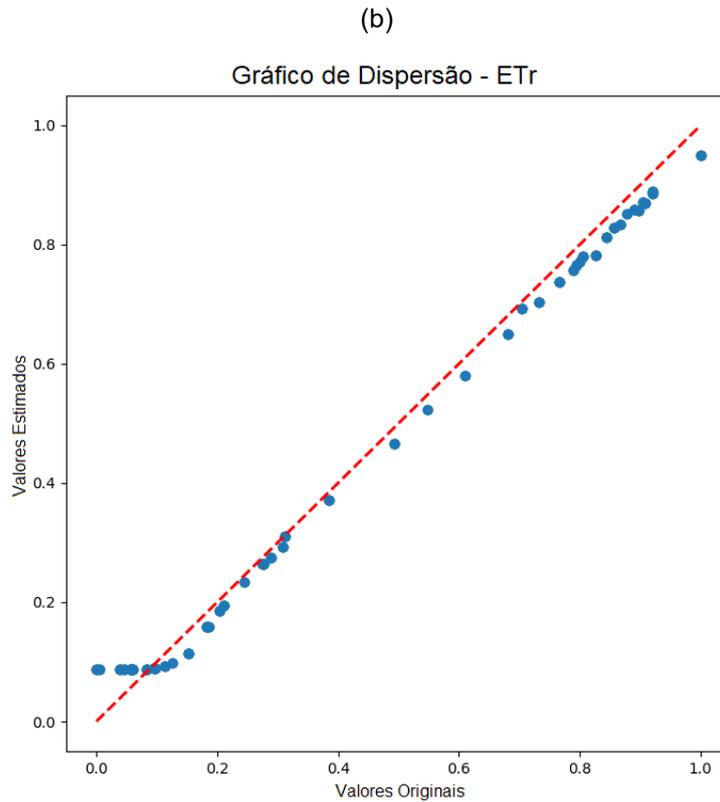
O modelo Random Forest apresentou bom desempenho na predição, apesar de ser inferior aos de DenseNet e FCN, sugerindo que a arquitetura montada com os atributos hiperparametrizados não foi suficiente, falhando em capturar plenamente a complexidade não linear, através das dependências temporais dos dados. Já a rede U-Net não foi eficaz para interpretar as características espaciais e temporais dos dados apresentados, onde ela

é mais indicada para tratar diretamente da segmentação de imagens e aprendizado espacial profundo. A rede neural profunda DenseNet ainda apresenta poucos trabalhos com modelagem de processos ambientais, mas com resultados robustos.

O modelo foi avaliado também quanto à sua precisão pelas análises de distribuição de erros (Figura 4a) e dispersão espacial dos pontos estimados (Figura 4b), onde o gráfico de resíduos mostra que a modelagem apresentou uma distribuição dos erros aproximadamente normal, voltada ao centro, em zero, com valores de média e mediana alinhadas, indicando ausência de viés sistemático, mostrando não existir ocorrência de superestimação ou subestimação dos valores estimados. Isto pode ser confrontado com a visualização da dispersão, que indica alta precisão preditiva, mostrando que o modelo DenseNet capturou adequadamente a relação entre as variáveis preditoras e a variável-alvo. As distorções observadas no gráfico de dispersão não influenciaram o resultado da estimação, confirmando a robustez do modelo adotado.

Figura 4 – Gráficos de Resíduos (a) e Dispersão (b) da variável-alvo (ETr).

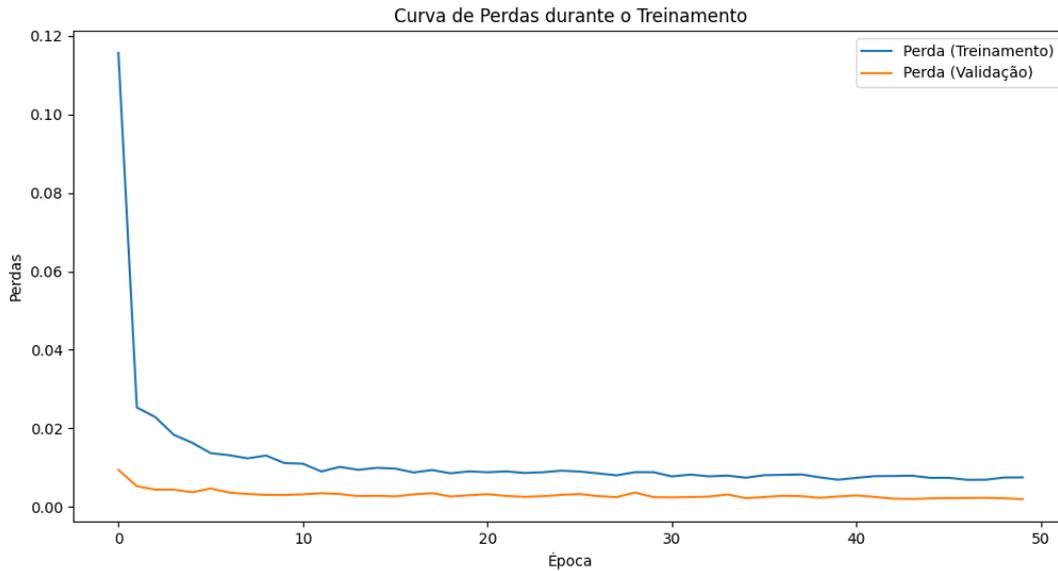




Fonte: Autor (2024)

Outra importante análise foi realizada durante a fase de treinamento do modelo DenseNet, onde as perdas de estimativas (Figura 5) indicam a qualidade do treinamento durante o progresso do aprendizado. A DenseNet apresentou consistência entre treinamento e validação, estabilizando a progressão após dez épocas, mostrando que rapidamente atingiu convergência, indicando ausência de overfitting, devido principalmente à evolução progressiva com proximidade entre as duas curvas, mais uma vez confirmando a capacidade de generalização dos dados.

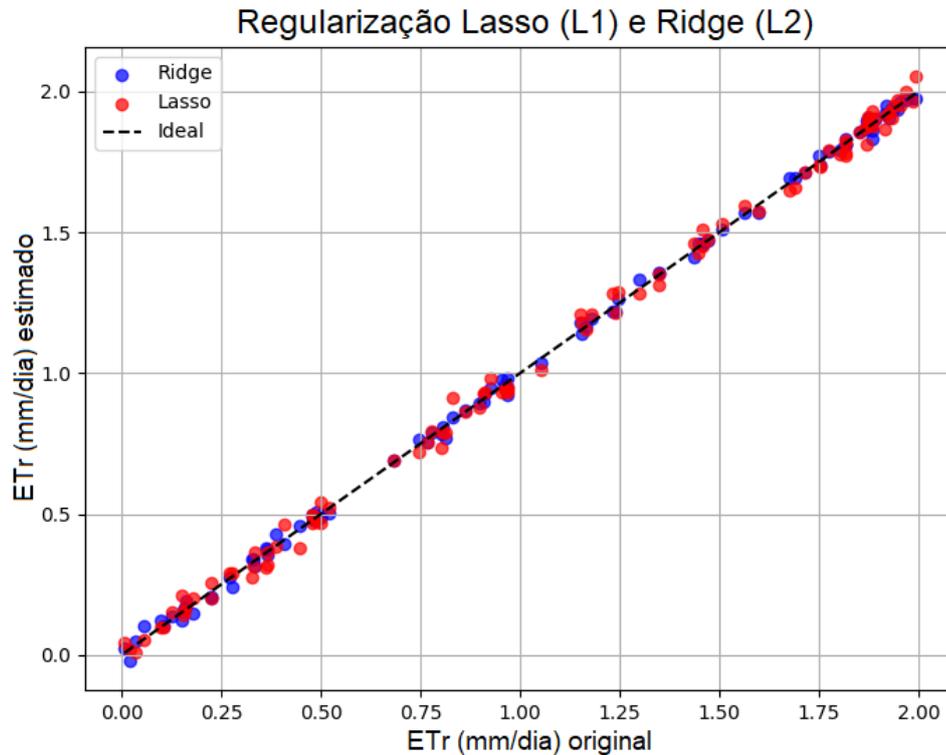
Figura 5 – Gráfico de Curvas de Perdas durante o treinamento do modelo DenseNet.



Fonte: Autor (2024)

Outra validação realizada na modelagem pelos modelos de IA foi o impacto da regularização Lasso (L1) e Ridge (L2) para estimação de ETr, garantindo melhoria da estabilidade e generalização. A implementação deste procedimento nos códigos construídos procura minorar a ocorrência de overfitting nos dados observados. No gráfico de dispersão apresentado (Figura 6) abaixo, o alinhamento entre os valores originais e estimados de ETr, onde os pontos representam as previsões do modelo DenseNet regularizado, demonstra distribuição equilibrada ao longo da linha ideal, tanto para a regularização Lasso (vermelho), quanto para a regularização Ridge (azul), ressaltando alta aderência aos valores reais, onde a penalização dos coeficientes auxiliou na melhoria da generalização.

Figura 6 – Gráficos de regularização L1 e L2 durante processamento dos modelos de IA.



Fonte: Autor (2024).

4.3.3 Análises dos resultados da predição com modelos de IA

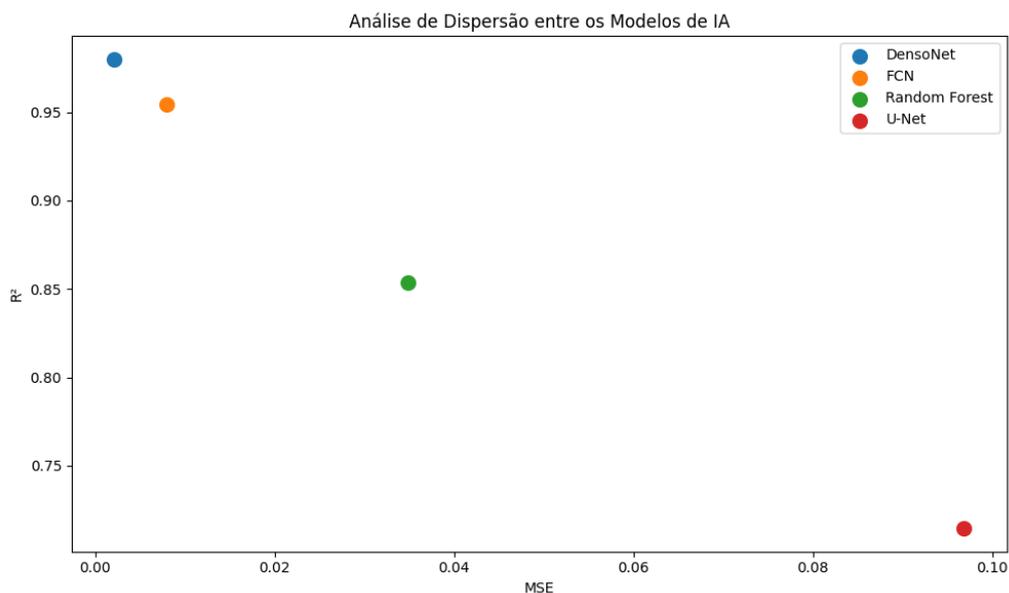
De forma geral, os modelos neurais profundos funcionam como caixas-pretas, ou seja, conseguem realizar predições com alta precisão, mas não fornecem uma explicação explícita dos processos subjacentes. Como consequência, há o risco de que incertezas indesejadas sejam incorporadas ao modelo, levando a erros sistemáticos que podem mascarar a complexidade espacial e não linear dos dados ambientais. A realização de análises diversas que expliquem os resultados obtidos é essencial para validar a qualidade da modelagem.

O algoritmo de DenseNet apresentou melhor acurácia dentre todos os modelos analisados, considerando como métricas de mse, rmse, mae, r^2 e cc (Tabela 2). Os resultados obtidos demonstram coerência com as análises exploratórias dos dados originais, os quais foram mantidos sem poda de *outliers*, com alta complexidade relacional, com vantagens na modelagem por alguns modelos. A padronização de procedimentos de

desenvolvimento dos modelos, como definição automática de hiperparâmetros, sem refinamentos, otimizados de forma automática, sem necessidade de ajustes sucessivos ou calibração manuais. Esse procedimento assegurou a consistência na generalização dos modelos, evitando ocorrência de erros fatais, como overfitting, ou outras inconsistências estruturais ou de subajustes.

Uma análise geral dos resultados (Figura 7) mostra que os algoritmos neurais profundos conseguiram maior capacidade preditiva para estimar os valores de evapotranspiração, frente à sua habilidade de capturar relações complexas e não lineares. A disponibilidade de dados balanceados e de menor espacialidade favoreceu as arquiteturas de redes profundas, reduzindo vieses e garantindo maior generalização dos modelos. DenseNet e FCN estão no quadrante ideal, com baixo valor de MSE e alto R^2 e CC. O algoritmo de RF conseguiu estimativa excelente, coerente com os padrões reportados na literatura, enquanto a U-Net não conseguiu ótima estimativa, sugerindo necessidades de ajustes nos parâmetros e/ou dificuldade em pontos com variabilidade extrema, indicando baixa adaptabilidade aos dados apresentados, além de necessidades de ajustes finos em seus parâmetros.

Figura 7 – Gráfico de análise de dispersão entre valores de MSE e R^2 para todos os modelos de IA avaliados.

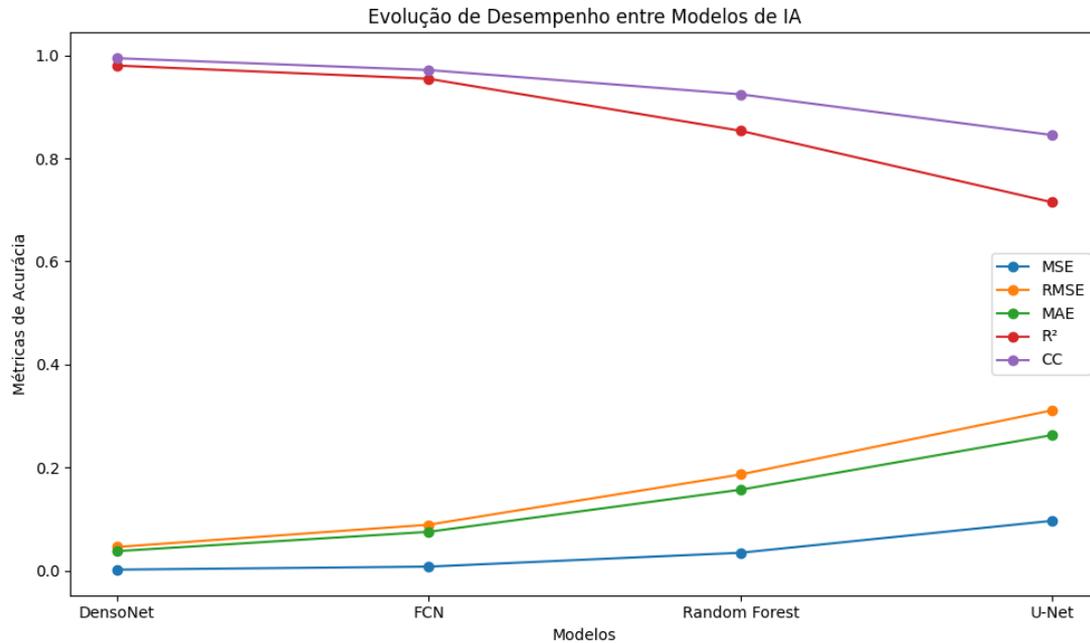


Fonte: Autor (2024).

Um gráfico comparativo entre as métricas normalizadas de MSE, RMSE, MAE, R^2 e CC (Figura 8) para os diferentes modelos também foi elaborado. A DenseNet manteve a melhor classificação em MSE, MAE e RMSE, refletindo um modelo com erros baixos e alta

capacidade preditiva, consistente com o R^2 calculado (0,98) e Coeficiente de Correlação CC (0,99).

Figura 8 – Gráfico comparativo de métricas normalizadas de desempenho entre modelos.



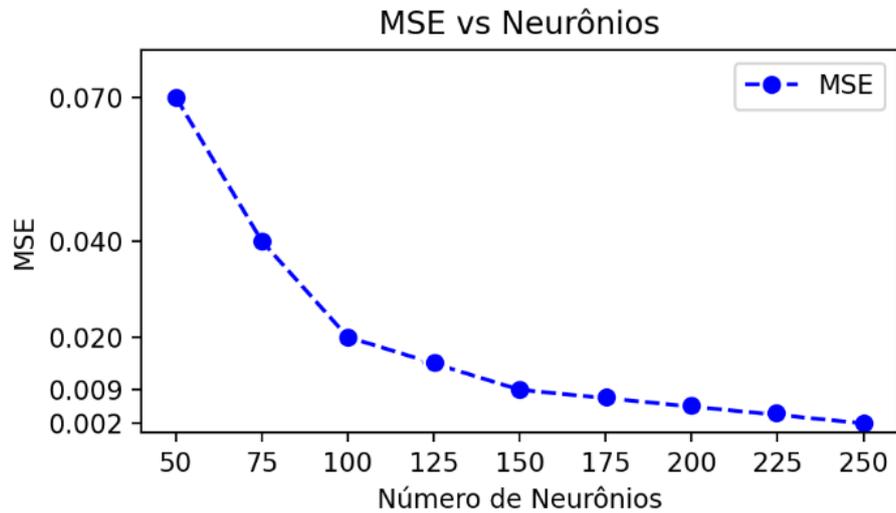
Fonte: Autor (2024)

O número de neurônios na camada oculta de um algoritmo neural, como o Dense, da biblioteca Keras® (Python), utilizada neste estudo para implementar a arquitetura DenseNet, é uma questão crítica, que deve ser controlada por mecanismos automáticos implementados para avaliar o seu desempenho preditivo. Um número inadequado de neurônios resulta em subajustes (underfitting), impactando a capacidade do modelo de capturar padrões complexos dos dados, ou resultar em sobreajuste (overfitting), permitindo que o modelo memorize padrões específicos avaliados na fase de treinamento, influenciando na sua capacidade de generalização. O DenseNet gerou os gráficos da evolução do erro quadrático médio (mse) versus número de neurônios na camada oculta (Figura 9a). O modelo atingiu valores menores progressivamente com o incremento de 157 neurônios, aproximadamente, quando os ganhos se tornaram marginais, indicando que o modelo poderia ser ajustado, onde a complexidade dos incrementos não contribuiu significativamente para a melhoria da predição. Para garantir esses resultados, mitigando riscos de sobreajuste, técnicas de validação cruzada, dropout, parada antecipada e regularização L1/L2 foram aplicadas. O modelo encerrou a estimativa de ETr com 250 neurônios na única camada escondida (Figura 9b), atingindo o melhor desempenho de

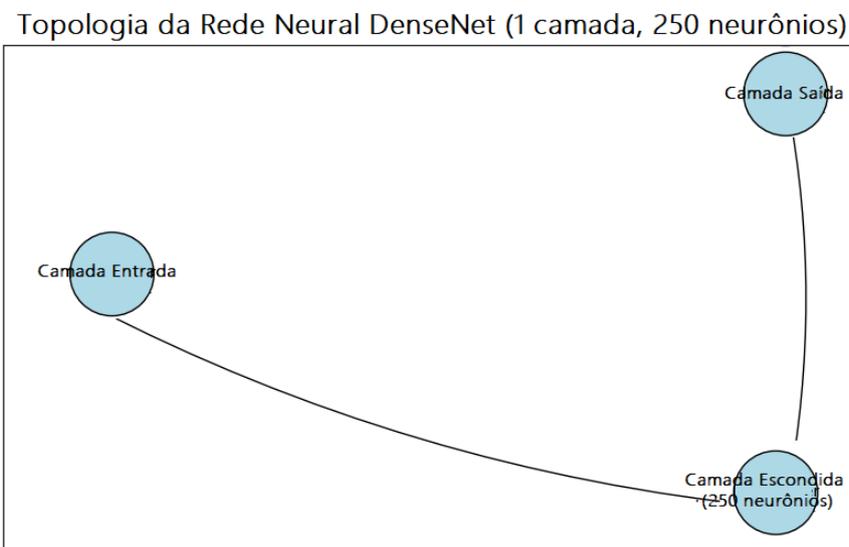
neurônios na camada oculta, resultando no melhor desempenho, reforçando sua capacidade de aprendizado e robustez para a modelagem de processos ambientais.

Figura 9 – Gráficos comparativo complementares de desempenho do modelo DenseNet.

(a)



(b)

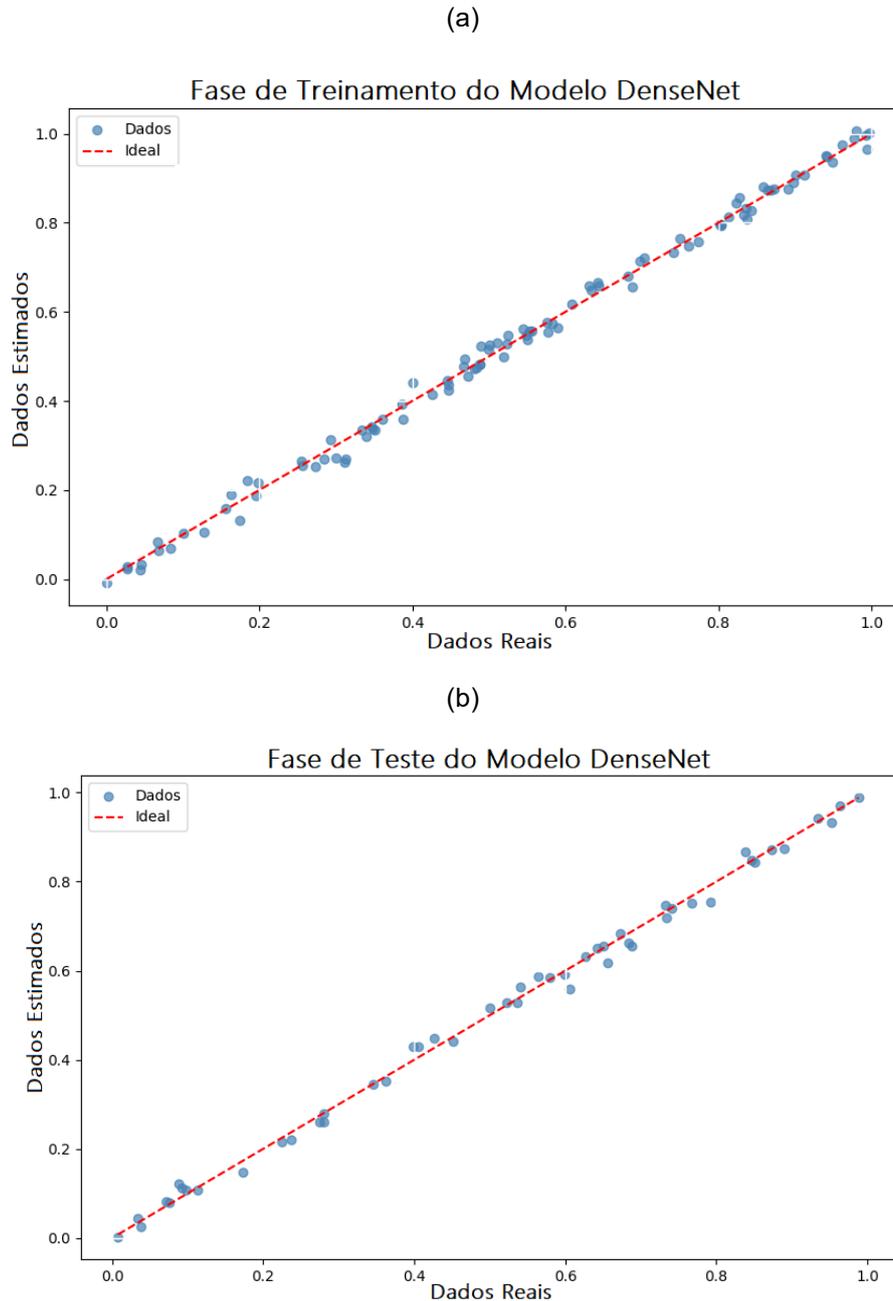


Fonte: Autor (2024)

O modelo DenseNet gerou gráficos da relação entre os dados reais e os dados estimados de ETr nas fases de treinamento e teste, visando avaliar seu desempenho. Na fase de treinamento (Figura 10a), conseguiu capturar corretamente as relações entre as

variáveis preditoras e alvo, com pequenas variações, confirmando erro residual mínimo, indicado anteriormente no gráfico de resíduos. Na fase de testes (Figura 10b), o modelo confirmou a robustez para a estimação da evapotranspiração real, sendo corroborada pelos valores de acurácia, demonstrando alta aderência entre os valores reais e estimados.

Figura 10 – Desempenho da DenseNet das fases de treinamento e teste.



Fonte: Autor (2024)

4.3.4 Análises explicativas da predição de ETr com o modelo DenseNet

Estudos recentes publicados na literatura, utilizando algoritmos de IA, têm ampliado o entendimento sobre os resultados obtidos, avaliando a influência das variáveis preditoras na variável-alvo, fornecendo entendimento mais claro sobre as relações envolvidas.

O modelo DenseNet utilizado neste trabalho gerou uma abordagem de PDP - *Partial Dependence Plots* (Kharate et al., 2024) para interpretar as relações entre ETr e NDVI de entrada com a ETr estimada. O gráfico de PDP para ETr (Figura 11a) evidenciou a relação negativa e aproximadamente linear entre ETr original e estimada. Isso indica que, à medida que a variável ETr aumenta, o impacto marginal sobre o modelo diminui, sugerindo saturação da resposta da DenseNet para essa variável.

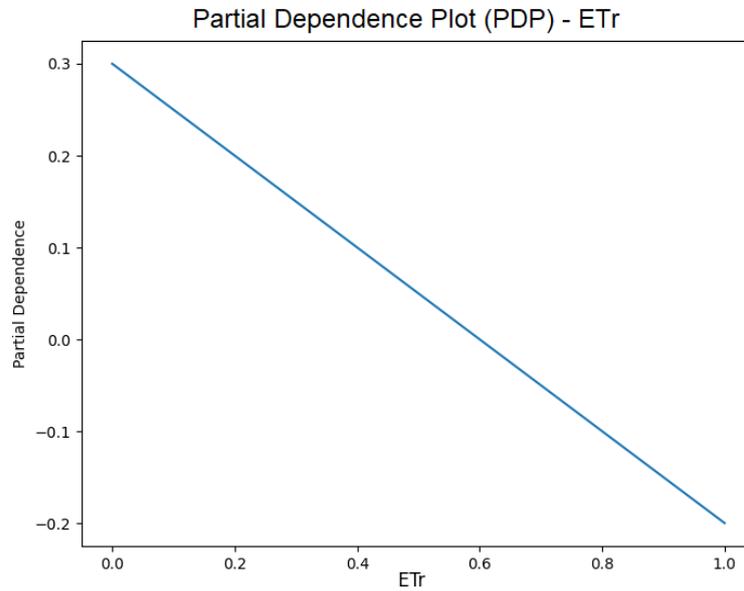
De maneira semelhante, o PDP para a variável NDVI (Figura 11b) evidenciou um comportamento linear e negativo, sugerindo que aumentos no índice de vegetação poderiam resultar em reduções na ETr estimada. Esse comportamento corrobora com a teoria fisiológica da vegetação na região de estudo, onde maior NDVI implica em maior cobertura vegetal e, conseqüentemente, menor ocorrência de evapotranspiração no solo exposto.

A não visualização de não-linearidade evidenciada nos gráficos de ETr e NDVI preditoras indica que os padrões de aprendizagem do modelo DenseNet ocorreram de forma majoritariamente linear, inexistindo efeitos contínuos, ou mesmo de interações complexas entre as variáveis preditoras e alvo, implicando na confiabilidade dos resultados obtidos de predição.

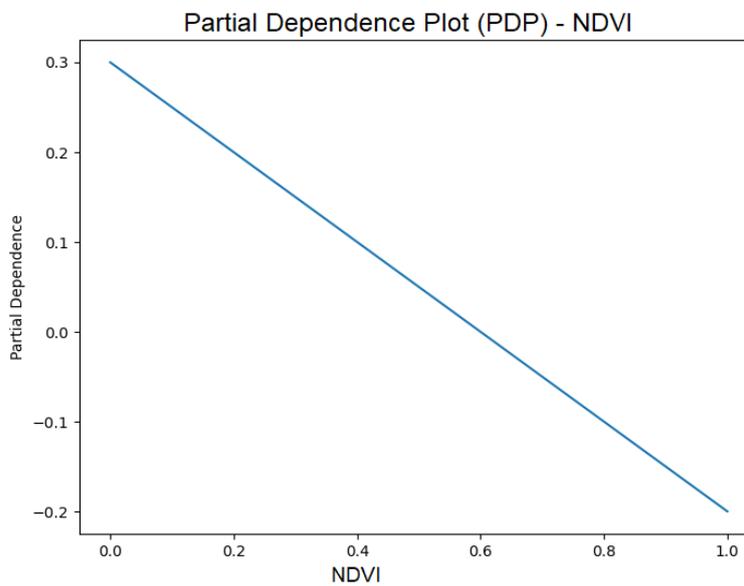
Assim, a linearidade visualizada nos gráficos sugere que as relações foram corretamente capturadas pelo modelo, sem irregularidades significativas ou impactos de pontos extremos.

Figura 11 – Análise da forma de relacionamento entre variáveis preditoras e alvo.

(a)



(b)



Fonte: Autor (2024)

Gráficos de importância de *features* usando SHAP - *Shapley Additive ExPlanation* (SALIH et al., 2024) foram implementados no modelo DenseNet, visando complementar a

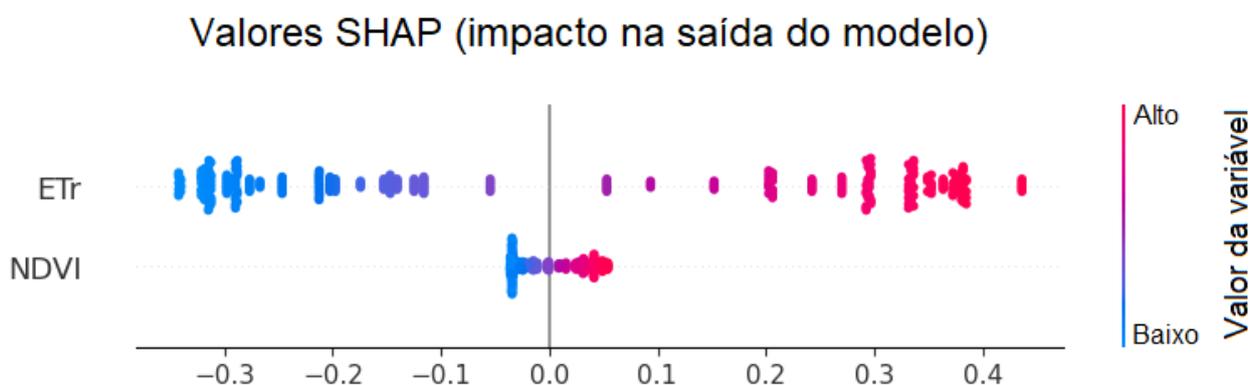
validação dos resultados produzidos. Os valores SHAP (Figura 12) para as variáveis preditoras, auxiliam no entendimento não apenas da direção de seus relacionamentos (positiva ou negativa), mas também da magnitude do impacto de cada característica, interpretando a contribuição de cada variável na previsão do modelo.

As variáveis foram ordenadas de acordo com o impacto médio, em valores absolutos, onde ETr preditora foi a mais influente na modelagem. O primeiro valor observado (0.4) indica o impacto médio da variável mais significativa, ressaltando que um aumento nesta variável tende a aumentar a previsão de ETr, implicando em uma relação direta e aproximadamente linear.

Com relação à NDVI, o impacto médio observado foi menor, mas ainda relevante, evidenciando uma relação fortemente negativa com a ETr estimada. Isso implica que os aumentos nos valores de NDVI tendem a reduzir a evapotranspiração real, o que pode ser sugerido pelo comportamento fisiológico da vegetação na área de estudo, em resposta à disponibilidade e uso eficiente de água.

Assim, os gráficos SHAP confirmam os resultados observados nos gráficos de PDP, indicando que ETr original foi a principal variável preditora do modelo, com NDVI desempenhando papel complementar, porém relevante, contribuindo para a modelagem dos valores estimados de ETr, aprimorando a predição e a relação entre as variáveis.

Figura 12 – Gráficos SHAP com direção, complexidade e relacionamento entre NDVI e ETr.



Fonte: Autor (2024)

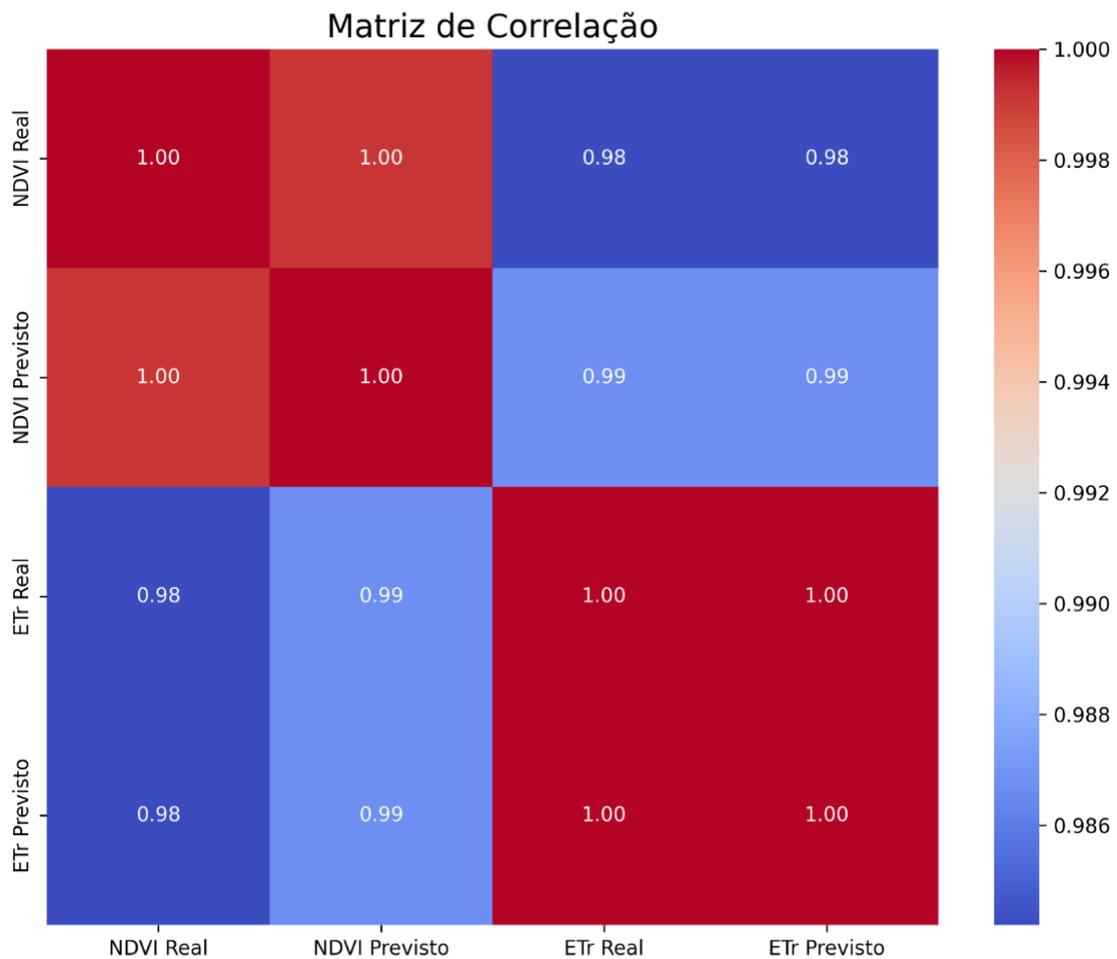
A matriz de correlação gerada pelo modelo DenseNet (Figura 13) reforça as relações evidenciadas entre as variáveis preditoras e alvo.

A consistência de dados modelados foi investigada quanto a discrepâncias ou padrões importantes, que corroborem análises anteriores. Foi observado que ETr real e ETr estimado mantêm correlação extremamente alta (1,00), o que evidencia eficácia na estimação da variável-alvo, confirmando as análises anteriores. A matriz também evidencia forte correlação entre o NDVI previsto e o ETr real (0,99), indicando haver relacionamento consistente entre estas variáveis, reafirmando a relação fisiológica prevista entre a vegetação e a evapotranspiração real.

A correlação um pouco inferior entre NDVI real e o ETr real (0,98) indica que a modelagem pode ter atenuado com pequenas variações em relação aos dados reais, priorizando previsões mais estáveis, confirmando correlação entre NDVI e ETr originais e ETr estimada foram capturadas pelo modelo. A ausência de correlações irregulares ou anômalas evidencia que o modelo não sofreu de problemas de sobreajuste, ou viés excessivo.

Assim, a matriz de correlação complementou as análises explicativas de PDP e SHAP, confirmando a robustez do modelo DenseNet na estimação de ETr, tendo como variáveis preditoras NDVI e ETr, com alta capacidade de generalização e precisão.

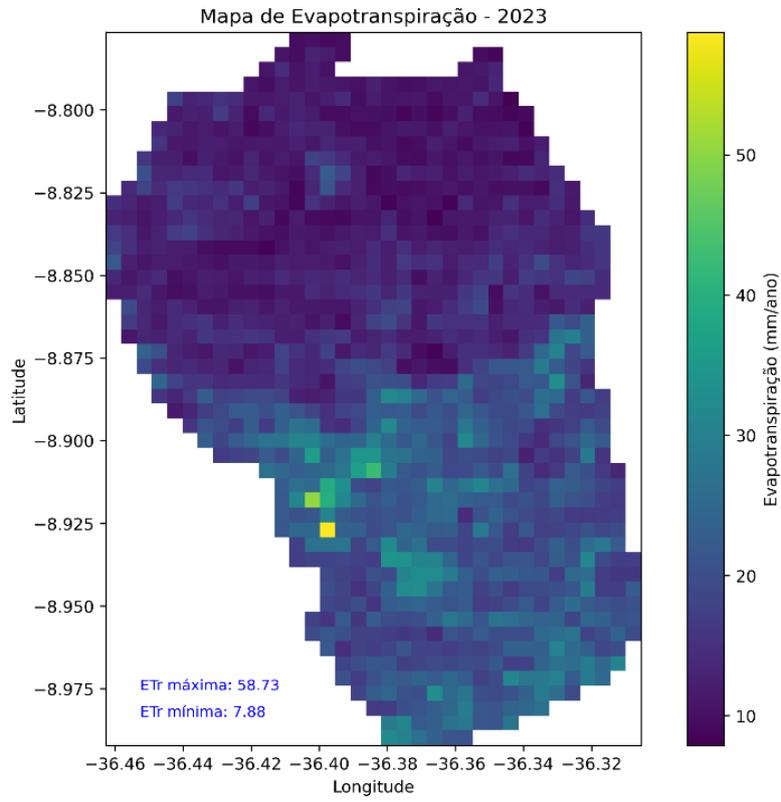
Figura 13 – Gráfico de correlação entre as variáveis predictoras e alvo do modelo de IA.



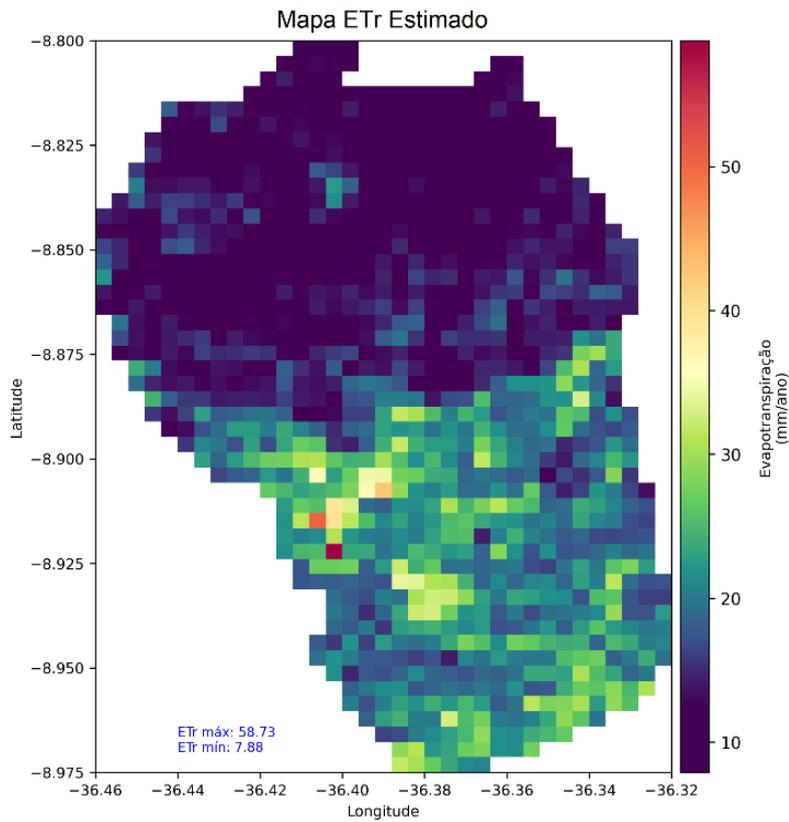
Fonte: Autor (2024)

Como produto resultante da modelagem ETr por modelos de inteligência artificial na área da pesquisa, é proposta a geração interativa de mapa de ETr estimada na região de estudo (Figura 14a), visando alimentação de sistemas diversos, como os de gerenciamento de recursos hídricos, ou de sistemas aplicativos de agricultura de precisão.

Figura 14 – Mapa interativo da área de estudos, pela modelagem da DenseNet, comparativamente aos dados reais.(a)



(b)



Fonte: Autor (2024)

4.3.4 CONCLUSÕES

Os resultados avaliados demonstraram a viabilidade da modelagem de ETr a partir de dados de NDVI e ETr oriundos de dados satelitais. A comparação entre o algoritmo de ML Random Forest e os três modelos de aprendizado de máquina profundo aplicados, evidencia a superioridade da ainda pouco investigada DenseNet para a estimação de variáveis ambientais, apresentando boa aproximação aos valores de NDVI e ETr, evidenciando sua capacidade em capturar as relações complexas envolvidas, ressaltada pelas análises estatísticas, onde a média e a assimetria foram preservadas, com pequena redução da variabilidade espacial, causada presumidamente, pela forte regularização preventiva aplicada na fase de treinamento do modelo.

As análises SHAP evidenciaram que NDVI e ETr apresentaram influência considerável sobre ETr estimada, com impacto linear negativo, corroborando a coerência da modelagem aos processos físicos envolvidos. A matriz de correlação ratificou a relação NDVI e ETr com ETr estimado, ressaltando que a DenseNet conseguiu excelente precisão tanto para as variáveis reais quanto para as previstas. Mecanismos robustos, como aplicação de validação cruzada, parada antecipada, regularização Lasso e Ridge, contribuíram para a estabilidade do modelo, minimizando riscos de sobreajustes ou outros vieses. A busca automática de atributos e parâmetros na construção dos modelos impactou positivamente para a generalização dos modelos.

Esta investigação atingiu seu propósito, mostrando que a abstração de complexidades pode gerar modelos robustos, embora sua aplicação seja limitada pela qualidade dos dados de entrada, que embora sejam representativos, possuem restrições de amostragem espacial e temporal, restringindo sua generalização para outras áreas de estudo ou períodos climáticos distintos.

REFERÊNCIAS

AL MAMUN, MD. A., SARKER, M. R., SARKAR, M. A. R., ROY, S. K., NIHAD, S. A. I., MCKENZIE, A. M., HOSSAIN, MD. I., & KABIR, MD. S. Identification of influential weather parameters and seasonal drought prediction in Bangladesh using machine learning algorithm. **Scientific Reports**, v. 14, n. 1, p. 566, 4 jan. 2024.

ALSENJAR, O., CETIN, M., AKSU, H. ET AL. Actual evapotranspiration estimation using metric model and landsat satellite images over an irrigated field in the eastern mediterranean region of turkey. **Med. Geosc. rev.** 5, 35–49 (2023). <https://doi.org/10.1007/s42990-023-00099-y>

BAFTI, A. G., AHMADI, A., ABBASI, A., KAMANGIR, H., JAMALI, S., & HASHEMI, H. Automated actual evapotranspiration estimation: Hybrid model of a novel attention based U-Net and metaheuristic optimization algorithms. **Atmospheric Research**, v. 297, p. 107107, jan. 2024.

BAŞAKIN, E. E.; EKMEKCIOĞLU, Ö.; ÖZGER, M. Providing a comprehensive understanding of missing data imputation processes in evapotranspiration-related research: a systematic literature review. **Hydrological Sciences Journal**, v. 68, n. 14, p. 2089–2104, 26 out. 2023.

BECK, H. E., ZIMMERMANN, N. E., MCVICAR, T. R., VERGOPOLAN, N., BERG, A., & WOOD, E. F. (2018). Present and future köppen-geiger climate classification maps at 1-km resolution. **Scientific Data**, 5(1), 180214. <https://doi.org/10.1038/sdata.2018.214>, 2018.

BEYER, M., AHMAD, R., YANG, B., & RODRÍGUEZ-BOCCA, P. Deep spatial-temporal graph modeling for efficient NDVI forecasting. **Smart Agricultural Technology**, v. 4, p. 100172, ago. 2023.

BURCHARD-LEVINE, V., BORRA-SERRANO, I., PEÑA, J. M., KUSTAS, W. P., GUERRA, J. G., DORADO, J., MESÍAS-RUIZ, G., HERREZUELO, M., MARY, B., MCKEE, L. M., DE CASTRO, A. I., SANCHEZ-ÉLEZ, S., & NIETO, H. Evaluating the precise grapevine water stress detection using unmanned aerial vehicles and evapotranspiration-based metrics. **Irrigation Science**, 29 abr. 2024.

BUTLER, R. A., PAPEŞ, M., VOGT, J. T., PAULSEN, D. J., CROWE, C., & TROUT FRYXELL, R. T. Human risk to tick encounters in the southeastern United States estimated with spatial distribution modeling. **PLOS Neglected Tropical Diseases**, v. 18, n. 2, p. e0011919, 14 fev. 2024.

CHAKHAR, A., ZITOUNA-CHEBBI, R., HERNÁNDEZ-LÓPEZ, D., BALLESTEROS, R., MAHJOUB, I., & MORENO, M. A. Assessing the Accuracy of Multiple Classification Algorithms Combining Sentinel-1 and Sentinel-2 for the Citrus Crop Classification and spatialization of the Actual Evapotranspiration Obtained from Flux Tower Eddy Covariance: Case Study of Cap Bon, Tunisia. **Proceedings of IAHS**, v. 385, p. 443–448, 19 abr. 2024.

DA SILVA, E. H. F. M., HOOGENBOOM, G., BOOTE, K. J., GONÇALVES, A. O., & MARIN, F. R. Predicting soybean evapotranspiration and crop water productivity for a tropical environment using the CSM-CROPGRO-Soybean model. **Agricultural and Forest Meteorology**, v. 323, p. 109075, ago. 2022.

DANDA, T. J., KUSANGAYA, S., MUSHORE, T. D., & MUJERE, N. Assessing the role of evapotranspiration in reducing surface temperatures in Harare using the SEBAL algorithm. **Urban Climate**, v. 49, p. 101442, maio 2023.

DAYAL, D., ET AL. Streamflow estimation using satellite-retrieved water fluxes and machine learning technique over monsoon-dominated catchments of India. **Hydrol. Sci. J.** 66 (4), 656–671, 2021.

DE SOUZA, C. A. A., JARDIM, A. M. D. R. F., DE SOUZA, L. S. B., JÚNIOR, G. D. N. A., ALVES, C. P., DE MORAIS, J. E. F., LOPES, D. D. C., STEIDLE NETO, A. J., SALVADOR, K. R. D. S., & DA SILVA, T. G. F. Intercomparison of micrometeorological variables, surface energy fluxes, and evapotranspiration in different landscapes of the Brazilian semi-arid region. **Agricultural and Forest Meteorology**, v. 341, p. 109679, out. 2023.

DE-VILLE, S., EDMONDSON, J., GREEN, D., STIRLING, R., DAWSON, R., & STOVIN, V. Effect of vegetation treatment and water stress on evapotranspiration in bioretention systems. **Water Research**, v. 252, p. 121182, mar. 2024.

DIYKH, M., ALI, M., JAMEI, M., ABDULLA, S., UDDIN, M. P., FAROOQUE, A. A., LABBAN, A. H., & ALABDALLY, H. Empirical curvelet transform based deep DenseNet model to predict NDVI using RGB drone imagery data. **Computers and Electronics in Agriculture**, v. 221, p. 108964, jun. 2024.

DOUNA, V., BARRAZA, V., GRINGS, F., HUETE, A., RESTREPO-COUBE, N., & BERINGER, J. Towards a remote sensing data based evapotranspiration estimation in Northern Australia using a simple random forest approach. **Journal of Arid Environments**, v. 191, p. 104513, ago. 2021.

ELZAIN, H. E., ABDALLA, O. A., ABDALLAH, M., AL-MAKTOUMI, A., ELTAYEB, M., & ABBA, S. I. Innovative approach for predicting daily reference evapotranspiration using improved shallow and deep learning models in a coastal region: A comparative study. **Journal of Environmental Management**, v. 354, p. 120246, mar. 2024.

FANG, Y., XU, L., CHEN, Y., ZHOU, W., WONG, A., & CLAUSI, D. A. A Bayesian Deep Image Prior Downscaling Approach for High-Resolution Soil Moisture Estimation. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 15, p. 4571–4582, 2022.

GEE. GOOGLE EARTH ENGINE, PLATFORM. **Produto GEE** Disponível em: <https://earthengine.google.com/>. Acesso em: 22 set. 2024.

GONG, H., CHENG, Q., JIN, H., & REN, Y. Effects of temporal, spatial, and elevational variation in bioclimatic indices on the NDVI of different vegetation types in Southwest China. **Ecological Indicators**, v. 154, p. 110499, out. 2023.

GONZÁLEZ PEREA, R., FERNÁNDEZ GARCÍA, I., CAMACHO POYATO, E., & RODRÍGUEZ DÍAZ, J. A. New memory-based hybrid model for middle-term water demand forecasting in irrigated areas. **Agricultural Water Management**, v. 284, p. 108367, jun. 2023.

GUO, H., HU, S., WANG, X., CHANG, M.-C., & LYU, S. Robust Attentive Deep Neural Network for Detecting GAN-Generated Faces. **IEEE Access**, v. 10, p. 32574–32583, 2022.

GUO, Y., ZHANG, L., HE, Y., CAO, S., LI, H., RAN, L., DING, Y., FILONCHYK, M. LSTM time series ndvi prediction method incorporating climate elements: a case study of yellow river basin. China. **Journal of Hydrology** 629, 130518, 2024. <https://doi.org/10.1016/j.jhydrol.2023.130518>, 2023.

HAO, P., DI, L., & GUO, L. Estimation of crop evapotranspiration from MODIS data by combining random forest and trapezoidal models. **Agricultural Water Management**, v. 259, p. 107249, jan. 2022.

KANG, W., TIAN, J., REEMT BOGENA, H., LAI, Y., XUE, D., & HE, C. Soil moisture observations and machine learning reveal preferential flow mechanisms in the Qilian Mountains. **Geoderma**, v. 438, p. 116626, out. 2023.

KHARATE, N., ANERAO, P., KULKARNI, A., & ABDULLAH, M. Explainable AI Techniques for Comprehensive Analysis of the Relationship between Process Parameters and Material Properties in FDM-Based 3D-Printed Biocomposites. **Journal of Manufacturing and Materials Processing**, v. 8, n. 4, p. 171, 6 ago. 2024.

KIM, S., LEE, E., HWANG, H.-T., PYO, J., YUN, D., BAEK, S.-S., & CHO, K. H. Spatiotemporal estimation of groundwater and surface water conditions by integrating deep learning and physics-based watershed models. **Water Research X**, v. 23, p. 100228, maio 2024.

KORKMAZ, M. SoC estimation of lithium-ion batteries based on machine learning techniques: A filtered approach. **Journal of Energy Storage**, v. 72, p. 108268, nov. 2023.

MA, Y., SUN, S., LI, C., ZHAO, J., LI, Z., & JIA, C. Estimation of regional actual evapotranspiration based on the improved sebal model. **Journal of Hydrology**, 619, 129283, 2023.

MEHMOOD R, BASHIR R, & GIRI KJ Deep generative models: a review. department of computer science, islamic university of science & technology, kashmir, jammu and kashmir, india. **Indian Journal of Science and Technology**, 16(7), 460–467, 2023.

MOD16A2.Earth Data, USGS, /Terra Net Evapotranspiration 8-day I4 **Produto Global** 500 m s/n grid. Disponível em: <https://lpdaac.usgs.gov/products/mod16a2v006/>. Acesso em: 19 set. 2024.

MOKHTARI, A., SADEGHI, M., AFRASIABIAN, Y., & YU, K. OPTRAM-ET: A novel approach to remote sensing of actual evapotranspiration applied to Sentinel-2 and Landsat-8 observations. **Remote Sensing of Environment**, v. 286, p. 113443, mar. 2023.

MOZAFFARI, H.; MOOSAVI, A. A.; NEMATOLLAHI, M. A. Predicting saturated and near-saturated hydraulic conductivity using artificial neural network and multiple linear regression in calcareous soils. **PLOS ONE**, v. 19, n. 1, p. e0296933, 10 jan. 2024.

MUÑOZ-CARPENA, R. et al. Convergence of mechanistic modeling and artificial intelligence in hydrologic science and engineering. **PLOS Water**, v. 2, n. 8, p. e0000059, 7 ago. 2023.

QGIS. **Produto QGIS Software**. Disponível em: <https://www.qgis.org/download/>. Acesso em: 25 set. 2024.

RAHIMPOUR, M., RAHIMZADEGAN, M. Assessment of surface energy balance algorithm for land and operational simplified surface energy balance algorithm over freshwater and saline water bodies in urmia lake basin. *Theor. Appl. Climatol.* 143 (3–4), 1457–1472, 2021.

REZAEI, M., MOUSAVI, S. R., RAHMANI, A., ZERAATPISHEH, M., RAHMATI, M., PAKPARVAR, M., JAHANDIDEH MAHJENABADI, V. A., SEUNTJENS, P., & CORNELIS, W. Incorporating machine learning models and remote sensing to assess the spatial distribution of saturated hydraulic conductivity in a light-textured soil. **Computers and Electronics in Agriculture**, v. 209, p. 107821, jun. 2023.

ROBERT, C. P.; ROUSSEAU, J. A special issue on Bayesian inference: challenges, perspectives and prospects. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 381, n. 2247, p., 15 de maio 2023.

SALIH, A., RAISI-ESTABRAGH, Z., GALAZZO, I. B., RADEVA, P., PETERSEN, S. E., MENEGAZ, G., & LEKADIR, K. A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME. **Advanced Intelligent Systems**, p. 2400304, 27 jun. 2024.

SENAY, G. B., PARRISH, G. E. L., SCHAUER, M., FRIEDRICH, M., KHAND, K., BOIKO, O., KAGONE, S., DITTMEIER, R., ARAB, S., & JI, L. Improving the Operational Simplified Surface Energy Balance Evapotranspiration Model Using the Forcing and Normalizing Operation. **Remote Sensing**, v. 15, n. 1, p. 260, 1 jan. 2023.

SHAN, W.; LIU, J.; ZHOU, J. Integrated method for intelligent structural design of steel frames based on optimization and machine learning algorithm. **Engineering Structures**, v. 284, p. 115980, jun. 2023.

SHREIM, H.; GIZZINI, A. K.; GHANDOUR, A. J. **Trainable Noise Model as an Explainable Artificial Intelligence Evaluation Method: Application on Sobol for Remote Sensing Image Segmentation**. ECRS 2023. **Anais**. Em: ECRS 2023. MDPI, 6 nov. 2023. Disponível em: <<https://www.mdpi.com/2673-4931/29/1/49>>. Acesso em: 12 dez. 2024

TARWIDI, D., PUDJAPRASETYA, S. R., ADYTIA, D., & APRI, M. An optimized XGBoost-based machine learning method for predicting wave run-up on a sloping beach. **MethodsX**, v. 10, p. 102119, 2023.

TURMAN, A. M., SOWBY, R. B., WILLIAMS, G. P., & HANSEN, N. C. Remote Sensing of Residential Landscape Irrigation in Weber County, Utah: Implications for Water Conservation, Image Analysis, and Drone Applications. **Sustainability**, v. 16, n. 21, p. 9356, 28 out. 2024.

VON KROGH, G.; ROBERSON, Q.; GRUBER, M. Recognizing and Utilizing Novel Research Opportunities with Artificial Intelligence. **Academy of Management Journal**, v. 66, n. 2, p. 367–373, abr. 2023.

XUE, H. et al. Runoff Estimation in the Upper Reaches of the Heihe River Using an LSTM Model with Remote Sensing Data. **Remote Sensing**, v. 14, n. 10, p. 2488, 23 maio 2022.

5 CAPÍTULO 2 – PREDIÇÃO DA CONDUTIVIDADE HIDRÁULICA SATURADA AVALIADA POR MODELOS DE INTELIGÊNCIA ARTIFICIAL

5.1 INTRODUÇÃO

O rápido desenvolvimento industrial e agrícola que ocorreu nas últimas décadas influenciou a expansão para áreas naturais e resultou em forte impactos nas áreas as terras exploradas, levando a uma maior preocupação e necessidade de se intensificar pesquisas sobre os solos e suas propriedades, principal destinatário de subprodutos gerados, também afetado por crescente contaminação residual por agroquímicos multissítios, como fertilizantes e pesticidas, que podem prejudicar suas propriedades, contaminando lençóis freáticos. Enriquecer o conhecimento de funcionamento hidrodinâmico do solo é fundamental, contribuindo para melhoria da produção de culturas, qualidade de águas que percolam sob raízes, preservando aquíferos subterrâneos. Na produção agrícola, estado de água no solo influencia desenvolvimento de sementes, crescimento de plantas, ciclo de nutrientes na zona radicular, teor de oxigênio no solo, facilitando trocas gasosas, contribuindo para a atividade biológica essencial da evolução de culturas, além da fluidez dos processos de infiltração no solo, e consequente escoamento superficial. No entanto, medições da condutividade hidráulica do solo, além de seus parâmetros de retenção de água nas camadas do solo demandam grande esforço de tempo e custos, especialmente em estudos de projeção regional, auxiliados por equipamentos e mão de obra especializados (BOUMA, 1989; VERECKEN, 1995; PACHEPSKY et al., 2006).

São habitualmente utilizados vários métodos experimentais para medir a condutividade hidráulica saturada (K_s) no terreno. Estes métodos incluem técnicas baseadas em laboratório e medições no terreno *in situ*. No entanto, tais abordagens têm os seus desafios e limitações, pois métodos laboratoriais dependem essencialmente da qualidade e precisão das amostras coletadas no solo, sob procedimentos padronizados de medições controladas. Sabe-se que estes métodos podem disponibilizar informações detalhadas das propriedades do solo, mas é comum não conseguirem capturar a variabilidade natural e a heterogeneidade das condições do terreno e seu entorno. Assim, os métodos de coleta em campo são aplicados em aplicações experimentais práticas. As medições de campo *in situ* incluem técnicas como os permeômetros de cabeça constante e de cabeça descendente, além do infiltrômetro de tensão. Estes métodos são frequentemente afetados por erros e imprecisões, que podem resultar em estimativa incorreta dos parâmetros específicos do solo, e de amortecimento tridimensional insuficiente do fluxo observado em bancada,

conduzindo a medições não confiáveis do K_s (Meter, 2024). O processo de medição pode ser complexo e confuso, particularmente quando se decide se se deve medir a condutividade hidráulica saturada ou não saturada, e ainda, qual o instrumento a utilizar para a aplicação (METER, 2024). As variações na composição e estrutura do solo têm um impacto significativo na exatidão destes métodos tradicionais. Por exemplo, os solos com estruturas heterogêneas ou com diferentes tamanhos de partículas podem conduzir a falsos valores de infiltração saturada, complicando o processo de medição (KEÏTA et al., 2023).

As investigações de campo mostraram que os valores reais de K_s podem divergir significativamente dos resultados laboratoriais devido a estes fatores específicos do solo (KEÏTA et al., 2023). As investigações de campo mostraram que os valores reais de K_s podem divergir significativamente dos resultados laboratoriais devido a estes fatores específicos do solo (KEÏTA et al., 2023). A partir de coleta de material em campo, modelos analíticos têm sido empregados para determinar a condutividade hidráulica saturada do solo, e adotam conceitos básicos do fluxo de água inerente ao processo, apresentando uma solução geralmente numérica de uma equação diferencial que rege o fenômeno.

Nas últimas décadas, tem-se pesquisado como alternativa estatística para mensuração de K_s , a aproximação indireta de propriedades hidráulicas do solo, utilizando funções de pedotransferência (FPT), recebendo aceitação considerável (TREJO-ALONSO et al., 2023; MOZAFFARI et al., 2024) na literatura acadêmica. As FPT foram introduzidas pela primeira vez a partir de equações de regressão empíricas, utilizando como recurso, informações que estão disponíveis no levantamento de solos, como teor de matéria orgânica, frações de areia, silte e argila, umidade volumétrica, além da densidade do solo. Aplicações de FPT reduzem custos operacionais e de tempo para estimação de K_s , sendo ideais para aplicação em grandes áreas, disponibilizando insumos para criação de mapas de condutividade hidráulica e análise de variabilidade espacial, essenciais para sistemas de gestão agrícolas e de recursos hídricos (GUPTA et al., 2021). Entretanto, a precisão de modelos com FPT depende da qualidade dos dados utilizados, da heterogeneidade do solo, assim como do alcance da representatividade espacial das variáveis selecionadas para refletir as propriedades da área de estudo. As FPT também podem não capturar outros processos físicos e químicos envolvidos na representação de K_s .

Recentemente, modelos numéricos suportados por sistemas computacionais têm sido empregados para capturar relações complexas não lineares e espaciais das variáveis envolvidas na estimação de K_s (FAÉ et al., 2020; MAHSA et al., 2020), conseguindo descrever a dinâmica da água no solo, relacionando a precisão com que os parâmetros físico-hídricos são determinados, incorporando, se modelados de uma maneira apropriada,

as incertezas intrínsecas a transferência de água no solo. A natureza imprecisa e qualitativa do conhecimento humano dificulta o tratamento matemático preciso exigido pelos modelos tradicionais. Por outro lado, a capacidade humana em manejar sistemas complexos, tratando incertezas, tem estimulado a busca por métodos de modelagem alternativos, que dêem relevância a esta categoria de informação. Existe a necessidade de se desenvolver modelos que forneçam de maneira mais eficiente e prática, aproximações que possam ser aplicadas, por exemplo, ao fluxo de água em três dimensões em condições de campo não saturado. As aproximações baseadas em técnicas de inteligência artificial (IA) podem satisfazer essa necessidade.

Publicações na literatura acadêmica apresentam técnicas de modelagem de K_s utilizando-se variados conjuntos de variáveis (ALBALASMEH et al., 2024; CHANG et al., 2024; DE RIJK et al., 2024), utilizando algoritmos de inteligência artificial (IA). Elas têm implicações positivas para práticas agrícolas atuais, além da gestão dos recursos ambientais. Modelagem com ferramentas de IA são empregadas principalmente para estimar valores de condutividade hidráulica saturada (K_s), contribuindo para otimizar atividades de irrigação. A previsão modelada do deslocamento dos corpos de água no solo auxilia na programação da irrigação, garantindo uso eficiente, racional e uniforme nos campos. Isso influencia no aumento da produtividade agrícola, maximizando conservação dos recursos hídricos disponíveis, fundamental em regiões com escassez de água (GUPTA et al., 2021). Além disso, condutividade hidráulica estimada por modelagem IA pode auxiliar em projetos de avaliação dos sistemas de irrigação, como aplicativos de agricultura de precisão (KARUNATHILAKE et al., 2023), permitindo programação a longo prazo de práticas agronômicas, com base nas previsões de IA associados a informações de séries temporais armazenadas, aumentando a capacidade de armazenamento de água no solo e água disponível.

A intensificação de previsões aprimoradas de K_s por modelagem IA possibilitará aplicá-la integrada em outros sistemas de gerenciamento de recursos climáticos, como cálculo de evapotranspiração e de parâmetros hidráulicos do solo, das propriedades físicas e do conteúdo de matéria orgânica, contribuindo para melhoria da estrutura funcional do solo, em aspectos como saúde e resistência do solo contra variações climáticas (CHOUDHARY et al., 2023; WANG et al., 2023). Pesquisas recentes com emprego de IA apresentam a flexibilidade de incorporação de variáveis ambientais, coletadas por técnicas de sensoriamento remoto (GONZALES-INCA et al., 2022; REZAEI et al., 2023), investigando sua contribuição para melhoria da acurácia na estimação de K_s , através da integração de técnicas de previsão climática baseadas em IA, com modelos híbridos (SLATER et al.,

2023), prevendo-se aumento significativo da exatidão e a robustez na sua predição. Estas técnicas permitem a assimilação de vastos dados climáticos, permitindo avaliar melhores previsões e maior precisão das propriedades do solo (MAHSA et al., 2020). Variáveis espaciais complexas, como variabilidade da precipitação, da temperatura e de outros fatores climáticos pouco estudadas em modelagem experimental, podem afetar a condutividade hidráulica do solo. Esta combinação de conhecimentos sobre o clima e modelos de IA cria uma compreensão mais dinâmica do ambiente do solo, aumentando o poder de previsão e apoiando melhores decisões de gestão agrícola e da água.

Um dos principais ganhos com a modelagem de Ks com algoritmos de aprendizado de máquina (ML), é a ampla capacidade de capturar relações complexas entre as variáveis preditoras utilizadas, além da disponibilidade de recursos para análise e interpretabilidade de resultados (MUÑOZ-CARPENA et al., 2023), como métodos de diagnóstico de modelos, tais como a análise da importância de modelos parciais e modelos substitutos, a partir de pequenos conjuntos de dados amostrais experimentais, garantindo sua acurácia para a previsão da condutividade hidráulica saturada.

A estimação de Ks combinando abordagens tradicionais com técnicas de IA baseadas em dados experimentais, aumenta a exatidão e a robustez das previsões, validando potencial de ambas as metodologias (MUÑOZ-CARPENA et al., 2023), dependendo da capacidade da IA para lidar com pequenos ou grandes conjuntos de dados complexos, melhorando a precisão das previsões hidrológicas. A padronização das etapas de desenvolvimento da aplicação de modelagem por inteligência artificial, incluindo fontes de dados, escolhas de algoritmos adequados, métodos avaliativos dos modelos, análises de resultados além de manter a transparência e o rigor científico dos resultados alcançados, aumentando a confiabilidade nos modelos computacionais (XIOURAS et al., 2022), contribuindo para o aprimoramento das pesquisas empregando modelagem computacional de processos ambientais com auxílio de ferramentas de IA.

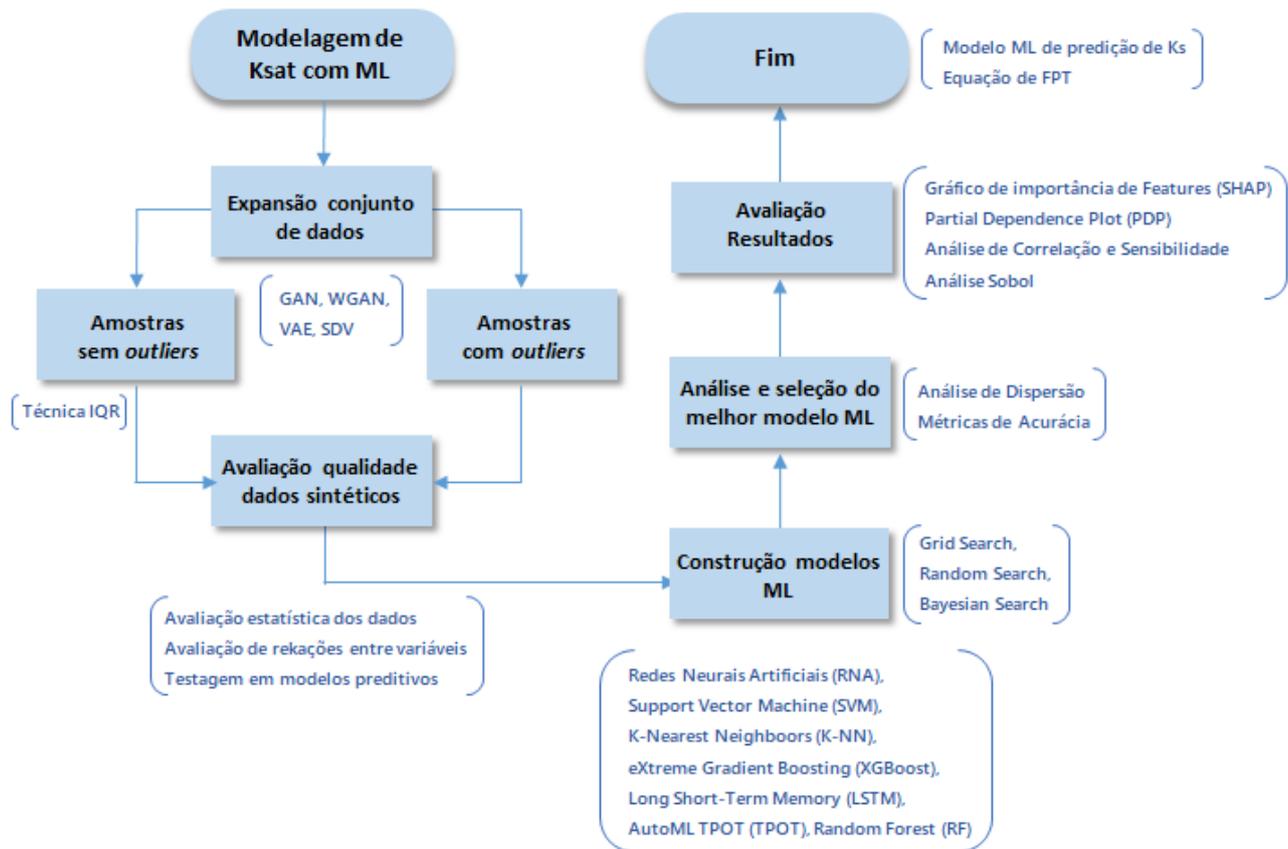
Sete algoritmos de IA foram aplicados para estimação da condutividade hidráulica saturada do solo utilizando pequena quantidade de dados experimentais disponíveis, que foram avaliados quanto às suas características, expansão e validação e sua qualidade por técnicas estatísticas poderosas (KUMAR et al., 2023), fornecendo método robusto de seleção do algoritmo de aprendizado de máquina (ML) adequado, a partir das características das variáveis preditoras e alvo envolvidas na simulação. Ao final, a modelagem de Ks pelo algoritmo selecionado, gerou proposta de um modelo de função de pedotransferência (FPT), baseada na melhor aproximação da condutividade hidráulica saturada, preservando ampla cobertura espacial do solo de estudo, viabilizando o potencial

das ferramentas de inteligência artificial para modelar propriedades hidrológicas e ambientais (ZHOU, 2024).

5.2 MATERIAIS E MÉTODOS

Experimento laboratorial foi realizado para estimar a condutividade hidráulica saturada de um solo, a partir da coleta técnica de material em campo. Cinco variáveis foram mensuradas para esta estimativa, sendo areia, silte, argila e densidade do solo, como variáveis de entrada, para uma variável-resposta (condutividade hidráulica saturada). Um conjunto de quarenta e quatro amostras foi avaliado por métodos estatísticos, procedendo-se poda de *outliers*. Dois conjuntos de dados amostrais (com poda e sem poda) foram então expandidos sinteticamente com GAN - Redes Adversariais Generativas (CHEN et al., 2021; VARASREE et al., 2024) e outras três técnicas de expansão de dados, avaliando-se qualidade e precisão relativamente aos dados experimentais originais por técnicas estatísticas e computacionais. Na sequência, sete modelos de inteligência artificial (aprendizado de máquina - ML), comumente descritos na literatura acadêmica para modelagem de processos ambientais, inclusive determinação de K_s , além da natureza dos dados, foram desenvolvidos e avaliados com estes conjuntos de dados expandidos, para modelagem na determinação de K_s . Visando garantir seleção do melhor modelo de ML para predição de resultados de modelagem, padronização de procedimentos para construção das arquiteturas foi aplicada, respeitando-se suas especificidades, consistindo em parametrização automática de seus atributos-chave (definição de hiperparâmetros) por métodos estatísticos de busca, seguido de avaliação de suas métricas de resultados. O modelo com melhores métricas de acurácia foi utilizado para predição da modelagem da condutividade hidráulica saturada do solo de estudo, além de gerar uma proposta de função de pedotransferência de maior representatividade espacial das variáveis preditoras do modelo (Figura 1).

Figura 1 – Descrição das etapas aplicadas para avaliação dos modelos de ML para estimação da condutividade hidráulica saturada (K_s). Fonte: Autores (2024).



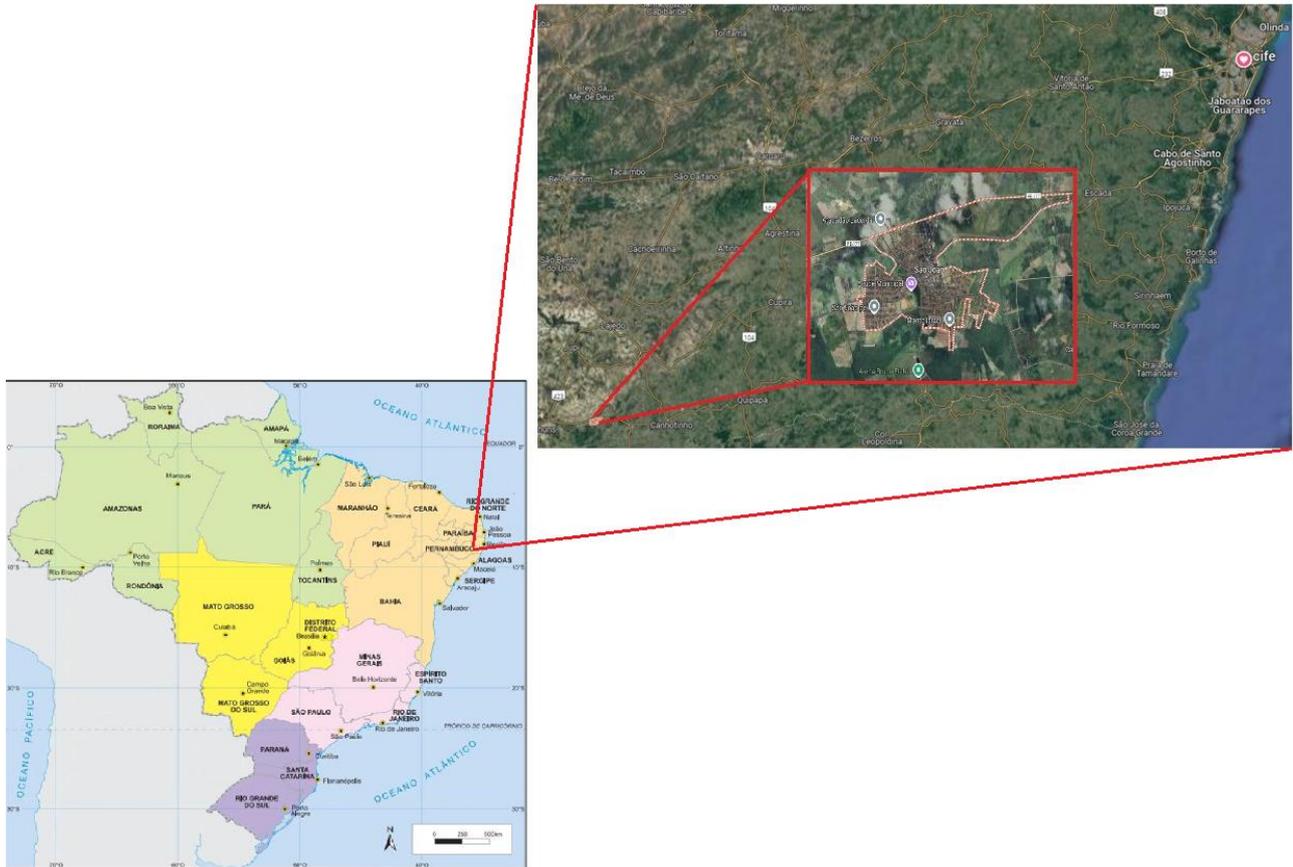
Fonte: Autor (2024)

5.2.1 Área de estudo, delineamento experimental e ensaios de campo

As propriedades hídricas do solo foram caracterizadas utilizando testes de campo de infiltração de água, realizados em área de origem agrícola, com área de aproximadamente 2.500 m², comumente utilizada para pastagem de gado, onde a vegetação cultivada é predominantemente de *Brachiaria decumbens* Stapf. A área fica localizada na fazenda Riacho do Papagaio, município de São João (Figura 2), na mesorregião do Agreste Meridional, no estado de Pernambuco, de coordenadas geográficas 8°52'30" de Latitude Sul, e 36°22'00" de Longitude Oeste, em altitude de 705 m. O clima observado é predominante o As', sendo equivalente a clima quente e úmido, conforme classificação de Köppen. O período chuvoso é previsto para o período de outono/inverno, com início em dezembro/janeiro e término até setembro. A precipitação média anual registrada para a

região é de aproximadamente 1.310 mm/ano. Aos dados hidrodinâmicos obtidos foram aplicados a estatística descritiva (máximo, mínimo, média, desvio-padrão e coeficiente de variação).

Figura 2 – Localização geográfica de São João/PE.



Fonte: Autor (2024)

5.2.2 Coleta de dados em campo e determinação das variáveis de estudo

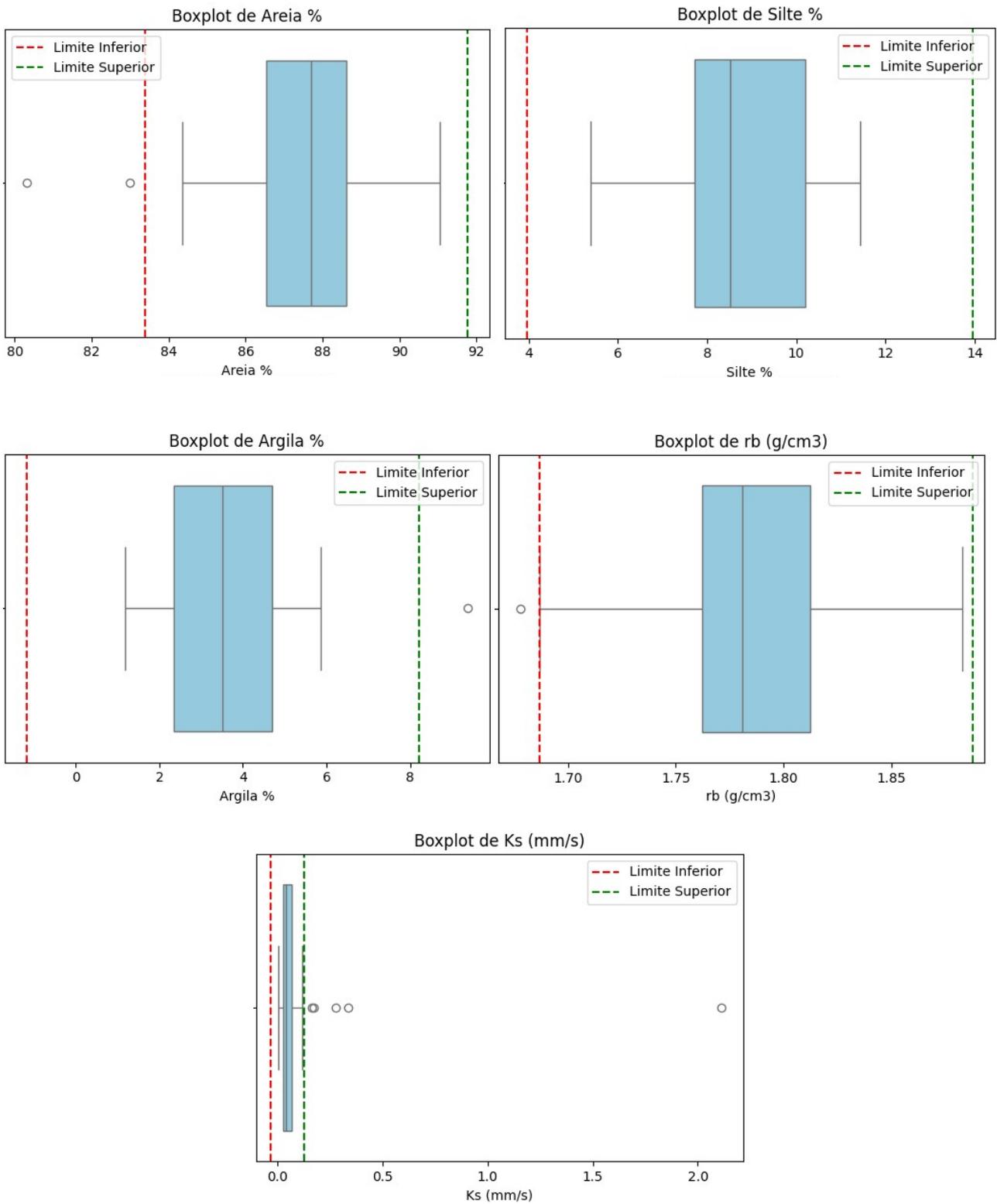
As amostras de solo coletadas no campo com um amostrador de Uhland adaptado, de capacidade para três cilindros de PVC, com 5 cm de altura, por 5 cm de diâmetro, precedidas da coleta de três amostras de solo com amostrador, utilizando-se cilindros de cinco centímetros de altura, sendo uma amostra representativa para a camada (0 - 5) cm, outra para (5 - 10) cm, e uma última para (10 - 15) cm. A coleta foi realizada a aproximadamente 1 m do eixo central do infiltrômetro, preservando o solo no local da infiltração. Todas as amostras foram acondicionadas apropriadamente, visando obtenção

da massa úmida M_u . Em seguida, levadas à estufa a 105°C por 24 horas, obtendo-se a massa seca M_s , para cálculo da umidade inicial Θ_0 , e da densidade do solo ρ_b . O método Beerkan (BRAUD et al., 2005; LASSABATÈRE et al., 2006) foi aplicado para estimativas dos parâmetros das curvas de retenção de água no solo $h(\theta)$ e da condutividade hidráulica $K(\theta)$, descritas, analiticamente, por cinco parâmetros: Dois de forma, m ou n e η , relacionados principalmente com a textura, e três de normalização θ_s , K_s e h_g , dependentes da estrutura do solo. Os parâmetros de forma foram obtidos a partir da curva de distribuição dos tamanhos das partículas $F(D)$ e da porosidade, enquanto os parâmetros de normalização foram determinados a partir de experimentos de infiltração. Para os ensaios de infiltração, foi utilizado um infiltrômetro de anel de quinze centímetros de diâmetro e aproximadamente oito centímetros de altura, encravado a um centímetro de profundidade, minimizando a perturbação estrutural do solo, visando assegurar infiltração axissimétrica na sua superfície (NETTO et al., 2013). Mais de sete volumes, variando de 70 a 250 mL de água foram infiltrados, atingindo estágio de regime de fluxo estacionário no local da infiltração (Souza et al., 2008), dependendo da taxa de infiltração, a fim de evitar uma carga hidráulica e consequente fluxo forçado no solo. Foram vertidos dentro do anel (infiltrômetro) volume de 100 mL de água, medindo-se o tempo necessário para a infiltração de cada volume. A partir da configuração geométrica, assumiu-se que a infiltração é tridimensional a uma pressão nula, ou seja, condição de umidade volumétrica saturada foi alcançada na superfície do solo (NETTO et al., 2013). Após a infiltração da água no solo, as amostras foram coletadas exatamente no centro do infiltrômetro utilizando um amostrador de Uhland modificado, com cilindros de plásticos de cinco centímetros de altura, para se obter o valor de umidade volumétrica final θ_f . Pesagem das amostras permitiram a obtenção da massa úmida M_u . Em seguida, as amostras foram levadas à estufa a 105°C , onde permaneceram 24 horas para a obtenção da massa seca M_s utilizada no cálculo da umidade volumétrica final por cinco repetições.

5.2.3 Tratamento de dados experimentais

O conjunto de 44 amostras experimentais, contendo 4 variáveis preditoras (areia, silte, argila e densidade do solo) e uma variável-alvo (condutividade hidráulica saturada) foi submetido a análise de *outliers*, utilizando-se a técnica de Interquartil - IQR (KORKMAZ, 2023), permitindo visualização de 9 pontos extremos (Figura 3), sendo areia = 2; silte = 0; argila = 1; $\rho_b = 1$; $K_s = 5$ ao ponto central.

Figura 3 – Box plot dos pontos extremos por variável, no conjunto amostral original.



Fone: Autor (2024)

Visando assegurar uma avaliação robusta do impacto da poda de outliers sobre a distribuição e a representatividade das relações entre variáveis, adotou-se uma abordagem integrada que combina Análise Exploratória de Dados (AED) (Tabela 1), avaliação de normalidade por QQ-plots, validação cruzada dos modelos preditivos e análise de importância de variáveis por SHAP. A AED foi conduzida seguindo os princípios clássicos de Tukey para inspecionar assimetria, caudas e pontos potencialmente influentes antes de qualquer intervenção, evitando decisões precipitadas de exclusão que possam distorcer o espaço amostral.

A Tabela 1 evidencia que as variáveis apresentam comportamentos estatísticos heterogêneos, típicos de dados ambientais. A areia exibe assimetria negativa (Skewness = $-1,24$), sugerindo concentração de massa à direita e cauda à esquerda; Silte tem assimetria próxima de zero ($0,07$), indicando distribuição quase simétrica; argila apresenta assimetria positiva moderada ($0,68$); a densidade do solo mostra valores próximos à simetria ($-0,07$) e curtose levemente platocúrtica ($-0,24$), sinalizando caudas relativamente curtas. Em contraste, Ks revela um padrão extremamente assimétrico (Skewness = $5,86$) e curtose muito elevada ($34,74$), compatível com caudas longas e presença de valores extremos raros, porém fisicamente plausíveis. Esses achados implicam que procedimentos de poda baseados apenas em limiares globais (por exemplo, regra do IQR de Tukey) podem remover informação substantiva, especialmente nas caudas direitas de Ks, com repercussões diretas na inferência e na aprendizagem de modelos.

A inspeção por QQ-plots (Figura 7 no seu documento e demais figuras anexadas nesta etapa) complementa a AED ao confrontar os quantis empíricos com os teóricos sob normalidade. Em areia e silte, observa-se bom alinhamento nos quantis centrais, com desvios discretos nas caudas, compatíveis com os valores de assimetria/curtose da Tabela 1. Em argila, a curvatura mais pronunciada no QQ-plot (concavidade para baixo em quantis baixos) confirma viés na cauda inferior; por sua vez, Ks afasta-se da reta de referência nos quantis superiores, ratificando cauda pesada. É oportuno lembrar que os QQ-plots são diagnósticos gráficos clássicos de normalidade/caudas — não testes formais — e seu uso deve ser articulado a procedimentos inferenciais como Shapiro–Wilk e Anderson–Darling, que oferecem poder estatístico diferenciado em amostras finitas (Shapiro–Wilk mais sensível a desvios no centro; Anderson–Darling mais sensível às caudas).

Diante desse panorama, a poda de outliers foi tratada como decisão de modelagem, não meramente de limpeza. Para Ks, cuja interpretação física sugere que valores extremos podem corresponder a micro-heterogeneidades do meio poroso (e não a erros de medida), optou-se por estratégias robustas: (i) transformações monotônicas (e.g., Yeo–

Johnson/Box-Cox) para reduzir assimetria antes da modelagem; (ii) validação cruzada com partições estratificadas por faixas de Ks para garantir que os extremos estejam presentes nos subconjuntos de treino e teste; (iii) emprego de perdas robustas e regularização no ajuste preditivo (evitando sobre-penalização de extremos). Em variáveis com caudas pouco informativas (e.g., outliers de areia devida a erro instrumental), considerou-se winsorização seletiva ou recalibração por quantis (*quantile mapping*) para preservar a ordem dos dados sem truncar relações. Essas decisões foram avaliadas antes e depois por QQ-plots e por testes formais de normalidade (Shapiro–Wilk e Anderson–Darling), permitindo aferir se a intervenção melhora a aderência sem empobrecer o suporte dos dados.

A validação cruzada foi empregada para quantificar o impacto da poda/transformação na generalização dos modelos. Independente do algoritmo, validou-se que a remoção indiscriminada de extremos de Ks reduz o desempenho em testes compostos por regimes de alta condutividade, refletindo quebra de suporte: o modelo deixa de “ver” padrões relevantes nas caudas e passa a extrapolar em produção. Esse efeito foi detectado tanto na métrica de acurácia preditiva em TSTR (treina-sintético / testa-real) — útil para avaliar dados gerados por modelos —, quanto na validação cruzada convencional sobre dados reais, reforçando que a cobertura de caudas é crucial para transferência de utilidade.

A análise de importância de variáveis via SHAP foi utilizada para interpretar o efeito das decisões de poda/transformação sobre o conjunto de relações do sistema. Em modelos com Ks como variável-resposta, os efeitos marginais e as interações entre textura (areia/argila/silte) e densidade emergem como determinantes (com sinais coerentes com a literatura de física do solo). Quando outliers de Ks são suprimidos de forma agressiva, observa-se rearranjo das contribuições SHAP: a importância atribuída a preditores associados a regimes extremos diminui, e o modelo passa a privilegiar tendências centrais — indício de viés de seleção induzido pela poda. Esse resultado justifica a preferência por tratamentos robustos e estratificação na validação, em vez de exclusão ampla.

Por fim, ao integrar (i) diagnósticos exploratórios (Tabela 1 + QQ-plots), (ii) testes formais de normalidade com ênfase nas caudas, (iii) validação cruzada e (iv) SHAP, obtém-se um protocolo reprodutível que equilibra controle de qualidade com preservação de estrutura. Tal protocolo é especialmente pertinente quando os dados reais são complementados por amostras sintéticas geradas por modelos (e.g., TabDDPM e CTGAN): nesses cenários, útil combinar TSTR/TRTR/TRTS na avaliação, pois a poda acrítica de outliers podem degradar a utilidade dominando um ganho aparente em fidelidade marginal.

Tabela 1 – Análise descritiva do conjunto amostral original

	Média	Mínimo	Máximo	Alcance	Assimetria	Curtose
Areia	87,50	80,31	91,04	10,73	-1,24	3,01
Silte	8,72	5,40	11,42	6,02	0,07	-0,75
Argila	3,80	1,17	9,38	8,21	0,68	1,91
Densidade do solo	1,78	1,68	1,88	0,21	-0,07	-0,24
Ks	0,11	0,00	2,11	2,10	5,86	34,74

Fonte: Autor (2024)

5.2.4 Expansão do conjunto de dados experimentais

Em trabalhos de pesquisas, é comum aplicar tratamento prévio sobre dados amostrais em procedimentos de expansão sintética, evitando-se problemas de *overfitting*, *underfitting*, viés, incapacidade para generalizar resultados (GUO et al., 2022; LIU et al., 2023). Entretanto, poucos trabalhos detalham como garantiram a preservação e replicação correta da distribuição e variabilidade espacial para o conjunto sintético (Mehmood et al., 2023). Outro importante fator diz respeito à manutenção das relações entre variáveis no conjunto amostral experimental, que podem prejudicar a acurácia dos modelos de ML que utilizam estes dados (MUÑOZ-CARPENA et al., 2023).

Embora a utilização de dados sintéticos esteja se tornando frequente em processos de modelagem de dados experimentais (LIU, 2023; KIRAN; KUMAR, 2024), gerados por diferentes métodos de expansão, faz-se necessário avaliar cuidadosamente os dados expandidos obtidos, de forma que não influenciem negativamente ou mascarem as performances dos modelos de ML aplicados, sendo importante seguir etapas avaliativas, como: (i) *geração dados sintéticos* (selecione método em função da natureza dos dados); (ii) *avaliação estatística dos dados* (média, variância, assimetria, curtose; Kullback-Leibler Divergence; Wasserstein Distance); (iii) *avaliação da variabilidade espacial* (Índice de Moran; Krigagem ou interpolação espacial); (iv) *avaliação de relações entre variáveis* (gráficos de dispersão entre variáveis; avaliação de matrizes de correlação); (v) *avaliação da diversidade* (número amostras únicas geradas; entropia de Shannon ou distância média ao ponto mais próximo); (vi) *teste em modelos preditivos* (desempenho em termos de MSE, RMSE, MAE, R² para regressão); (vii) *validação cruzada* (combinação de dados reais e sintéticos para validar o impacto predição nos modelos de IA). A escolha do método de expansão deve ser baseada em alguns fatores-chave, como: Complexidade dos dados

originais; no objetivo da modelagem; na disponibilidade de dados; requisitos de desempenho, e facilidade de implementação (JEBRAEELI et al., 2024; ZHOU et al., 2024).

Dada a pequena cardinalidade amostral e a variabilidade espacial limitada das medições originais, empregou-se expansão artificial do conjunto de dados com geradores de natureza diversa, visando viabilizar e estabilizar o ajuste dos sete modelos de IA selecionados. Foram considerados modelos generativos clássicos e recentes para dados tabulares: Redes Adversárias Generativas (GAN), incluindo a formulação Wasserstein (WGAN); Autoencoders Variacionais (VAE); a família Synthetic Data Vault (SDV) — notadamente CTGAN e GaussianCopula, e, como método de ponta, o TabDDPM (difusão para dados tabulares). Essa diversidade buscou cobrir regimes de geração adversarial, variacional, copular e difusivo, que diferem em como tratam caudas, modos e dependências entre variáveis.

Para garantir a qualidade estatística e a utilidade prática dos dados sintéticos, definiu-se um protocolo avaliativo em múltiplas camadas, aplicado antes e depois da expansão: (i) Testes de aderência univariados (Tabela 6) — Kolmogorov-Smirnov, Anderson-Darling e Shapiro-Wilk — para aferir similaridade marginal entre distribuições real e sintética, com ênfase em caudas (AD) e centro (SW); (ii) Análise descritiva dos conjuntos expandidos (Tabela 2), cotejando média, amplitude, assimetria e curtose com os valores do conjunto original (Tabela 1), a fim de avaliar preservação de escala e forma; (iii) Validação com modelos de ML (Tabela 3) com validação cruzada K-Fold, em métricas MAE, MSE, RMSE e R^2 ; (iv) avaliação explícita de utilidade por transferência usando TSTR/TRTR/TRTS e curvas de aprendizado (learning curves) em função do número de amostras sintéticas, com IC95% por bootstrap, além de métricas de distância (e.g., Wasserstein) e diagnósticos gráficos (histogramas, QQ-plots, ECDF e KDE-2D). Esse protocolo impede que ganhos aparentes em fidelidade marginal (univariada) ocultem perdas em utilidade na tarefa de interesse.

Os resultados mostraram diferenças sistemáticas entre métodos. TabDDPM obteve os melhores valores de TSTR R^2 e curvas de aprendizado mais favoráveis, apesar de vieses localizados (p. ex., Argila com excesso de massa na cauda inferior e Ks com encolhimento de caudas). A GaussianCopula destacou-se na fidelidade marginal (menores distâncias univariadas), mas com utilidade inferior na tarefa preditiva. GAN/WGAN apresentaram instabilidade (oscilação das perdas e indícios de colapso de modos), refletindo-se em R^2 negativos em TSTR e desempenho inconsistente; VAE gerou distribuições suavizadas com sub-dispersão (amplitude reduzida), o que penalizou a aprendizagem em regimes extremos.

A Tabela 2 (conjunto sem poda de *outliers*, expandido por TabDDPM) evidencia o efeito típico da difusão: redução de assimetria/curtose e compactação das distribuições, com melhor alinhamento aos quantis centrais — mas com perda parcial de extremos. Em números: Ks passou de Skewness 5,86 e Curtose 34,74 (Tabela 1) para 0,97 e 1,06 (Tabela 2), respectivamente; Areia reduziu a assimetria de $-1,24$ para $-0,13$, e a curtose de 3,01 para $-0,23$; Argila inverteu o sinal da assimetria ($0,68 \rightarrow -0,42$) e reduziu a curtose ($1,91 \rightarrow -0,95$); Silte manteve leve platocurtose; e a Densidade do solo permaneceu próxima da simetria, com discreto achatamento. As figuras (histogramas, QQ-plots e *scatters*) confirmam: rb e Areia/Silte com boa aderência no centro; Ks com cauda direita sub-representada; e Argila com acúmulo na faixa inferior, demandando calibração por quantis e/ou condicionamento por faixas. Esses padrões explicam a não-monotonicidade observada na learning curve de TSTR: ao inflar n sem calibrar as caudas, a transferência degrada; quando n cresce ainda mais, a cobertura do espaço melhora e o desempenho se recupera.

No tocante à validação preditiva (Tabela 3), a expansão por TabDDPM sem poda foi a que melhor preservou a generalização dos sete modelos testados, sobretudo quando combinada com pré-processamento monotônico (p. ex., Yeo-Johnson) e pós-ajuste por quantis nas variáveis críticas (Argila e Ks). Em contraste, a expansão por GAN/WGAN — mesmo com ajustes de *learning rate*, *batch size* e *TTUR* — teve dificuldade em capturar correlações tabulares sem incorrer em colapso de modos, cenário conhecido na literatura quando a dimensionalidade efetiva é alta e a amostra é pequena. VAE tendeu a encurtar amplitudes e suavizar caudas; CTGAN (SDV) mitigou parcialmente o problema, porém não superou o TabDDPM em TSTR, embora mantenha interesse por facilidade de uso e boa fidelidade univariada.

Importante destacar que o conjunto com poda de *outliers* não apresentou resultados satisfatórios em nenhum dos métodos. A remoção agressiva de extremos de Ks, em particular, quebrou o suporte do problema: os modelos passaram a extrapolar em regimes de alta condutividade e perderam importância relativa de preditores associados a esses regimes (confirmado pela análise SHAP), o que se refletiu em pior desempenho em validação cruzada e TSTR. Assim, a estratégia recomendada é evitar a poda e, em seu lugar, aplicar tratamentos robustos: transformações monotônicas por variável, condicionamento por faixas (especialmente para Argila), reponderação de caudas em mini-batches (para Ks) e pós-ajuste por quantile mapping nas bandas 1–5% e 95–99%. Esses ajustes preservam extremos fisicamente plausíveis e estabilizam a transferência de utilidade.

Tabela 2 – Análise descritiva do conjunto amostral sem poda de pontos extremos, expandido por GAN.

	Média	Mínimo	Máximo	Alcance	Assimetria	Curtose
Areia	87,71	84,35	91,04	6,70	-0,13	-0,23
Silte	8,71	5,40	11,42	6,02	0,00	-0,78
Argila	3,58	1,17	5,86	4,69	-0,42	-0,95
Densidade do solo	1,78	1,72	1,85	0,14	0,14	-0,68
Ks	0,04	0,00	0,12	0,11	0,97	1,06

Fonte: Autor (2024)

Assim, embora métodos adversariais (GAN/WGAN) e variacionais (VAE) constituam *baselines* relevantes, a difusão tabular (TabDDPM) é, neste estudo, a opção preferencial para ampliação sem poda, desde que acompanhada dos ajustes direcionados descritos. A GaussianCopula/CTGAN pode ser mantida como complemento quando a prioridade for fidelidade marginal por variável, compondo um portfólio híbrido de expansão.

5.2.5 Construção dos modelos de aprendizado de máquina (ML)

Sete modelos de inteligência artificial foram selecionados, tendo como premissa, pesquisas na literatura sobre modelagem de processos ambientais utilizando algoritmos de aprendizado de máquina, além das características dos dados experimentais (MUÑOZ-CARPENA et al., 2023). Os algoritmos de TPOT (WEN; LI, 2022), Redes Neurais Artificiais (MOZAFFARI et al., 2024), Random Forest (KANG et al., 2023; REZAEI et al., 2023), K-Nearest Neighbors (HE et al., 2023; VELOSO et al., 2022), Support Vector Machine (SINGH et al., 2022; FAOUZI et al., 2023), LSTM (LÜ et al., 2024), XGBoost (LI et al., 2023; AHMADISHARAF et al., 2024) foram desenvolvidos para a modelagem do processo de condutividade hidráulica saturada, a partir dos dois conjuntos de dados expandidos (com e sem poda), buscando capturar as relações entre as variáveis de entrada e saída, utilizando *frameworks* apropriados para cada algoritmo de ML na linguagem de programação Python®.

Visando facilitar a robustez de cada modelo construído, assim como possibilidade de comparação de resultados, todos utilizaram estratégias similares, com etapas claras de: leitura de dados: pré-processamento (normalização com função *StandardScaler*, divisão de

dados com 30% dos dados para teste); treinamento (RNA com função ativação relu, solver adam, parada antecipada), avaliação (métricas MAE, MSE, RMSE e R^2 , variância explicada e MAPE); Otimização de hiperparâmetros (seleção de métodos de buscas Grid Search, Random Search, Bayesian Search, RNA para tamanho camadas ocultas, taxa de aprendizado, e regularização α); avaliação de robustez (validação cruzada com 5 divisões); avaliação de resultados (análise de resíduos com histogramas e gráficos de dispersão, explicabilidade com PDP e SHAP, análise de sensibilidade Sobol); avaliação de resultados.

5.2.6 Hiperparâmetros dos modelos de ML por mecanismos de busca

Três métodos de buscas usualmente utilizados na literatura, Grid Search (TARWIDI et al., 2023), Random Search (SHAN et al., 2023), e Bayesian Search (ROBERT; ROUSSEAU, 2023) foram aplicados em todas as arquiteturas desenvolvidas, visando definição automática do melhor conjunto otimizado de hiperparâmetros. Cada modelo de IA foi elaborado para utilizar hiperparâmetros apropriados às suas características funcionais (VON KROGH et al., 2023). Dentre os principais atributos avaliados em cada modelo de ML, destacam-se: Redes Neurais Artificiais (*hidden_layer_sizes*, *activation*, *alpha*, dentre outros), Random Forest (*n_estimators*, *max_depth*, *min_samples_split*, e outros), XGBoost (*n_estimators*, *max_depth*, *learning_rate*, *subsample*), TPOT (*configurações automáticas com busca genética*), SVM (*C*, *epsilon*, *kernel*, etc.), LSTM (*units*, *activation*, com adaptação ao formato de entrada temporal), e KNN (*n_neighbors*, *weights*, *p*). O método de pesquisa de atributos *Bayesian Search* foi o de melhor desempenho geral, comparativamente a *Random Search* e *Grid Search*, nos sete modelos testados, com os menores valores de MSE e RMSE e maior valor de R^2 .

5.2.7 Desempenho dos Modelos de ML

As métricas utilizadas estimaram MAE, MSE, RMSE e R^2 . Análises de estatística realizados por meio de teste ANOVA (Kruskal-Wallis) não indicou diferença significativa entre os métodos, sugerindo que as abordagens obtiveram resultados comparáveis. O melhor desempenho foi alcançado pelo modelo de RNA, seguido de LSTM e TPOT (Tabela 3).

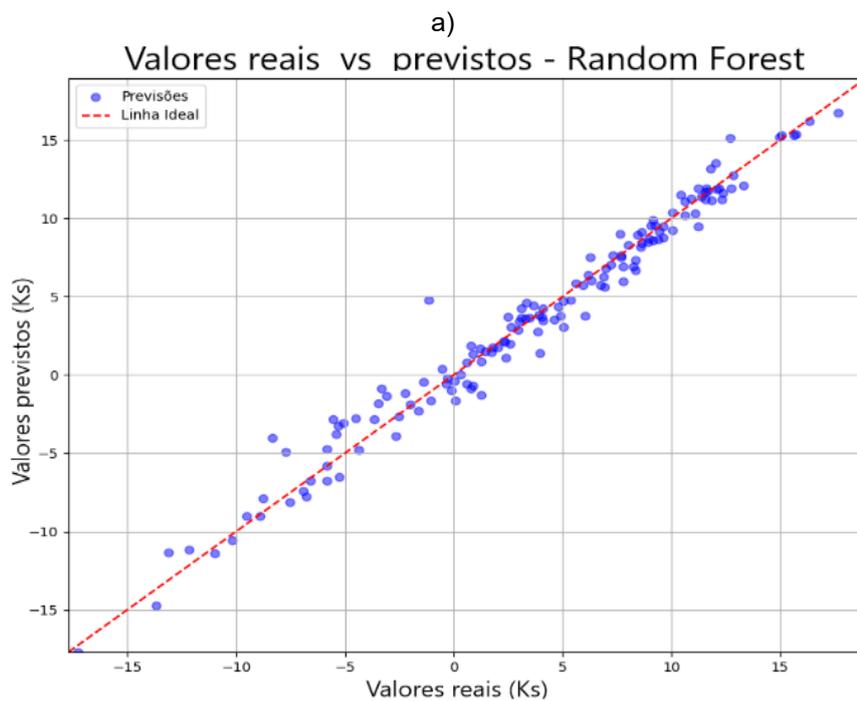
Tabela 3 – Classificação por desempenho dos modelos de IA treinados com conjunto de dados sintéticos.

	Método Busca	MSE	RMSE	MAE	R²
RNA	Bayesian Search	0,01	0,09	0,08	0,99
LSTM	Bayesian Search	0,07	0,27	0,09	0,99
TPOT	Bayesian Search	0,01	0,12	0,10	0,98
XGBoost	Bayesian Search	0,02	0,13	0,10	0,98
RF	Bayesian Search	0,09	0,30	0,10	1,00
Knn	Grid Search	0,03	0,16	0,11	0,97
SVM	Random Search	0,09	0,30	0,30	0,99

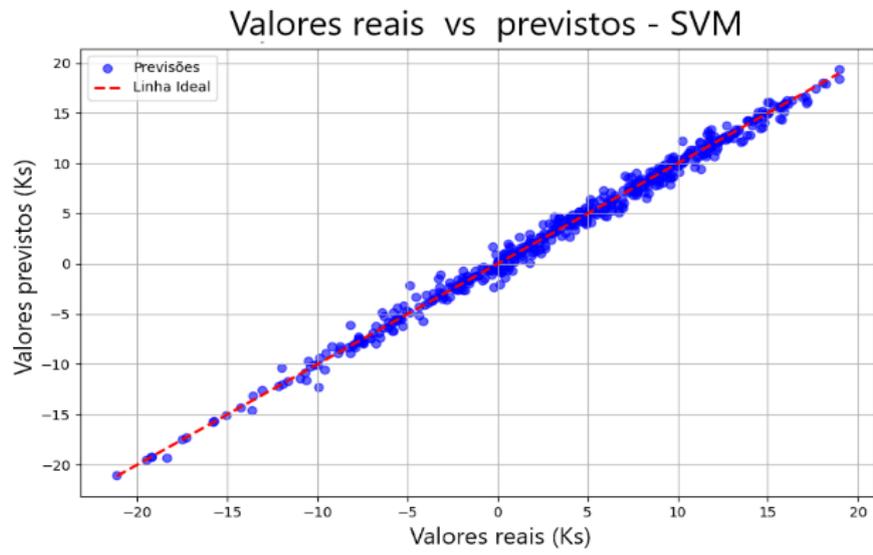
Fonte: Autor (2024)

Os resultados obtidos pelos sete modelos de IA foram próximos, do ponto de vista de valores de desempenho (Figura 4).

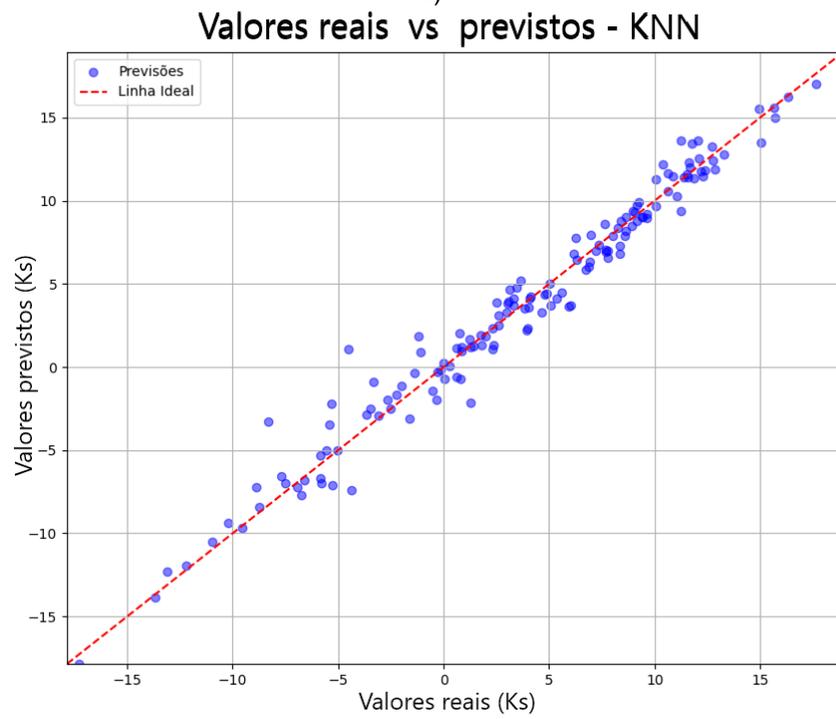
Figura 4 – Gráficos de desempenho (valores reais versus previstos) de Ks pelos modelos de IA: a) Random Forest; b) SVM; c) KNN; d) XGBoost; e) TPOT; f) RNA.



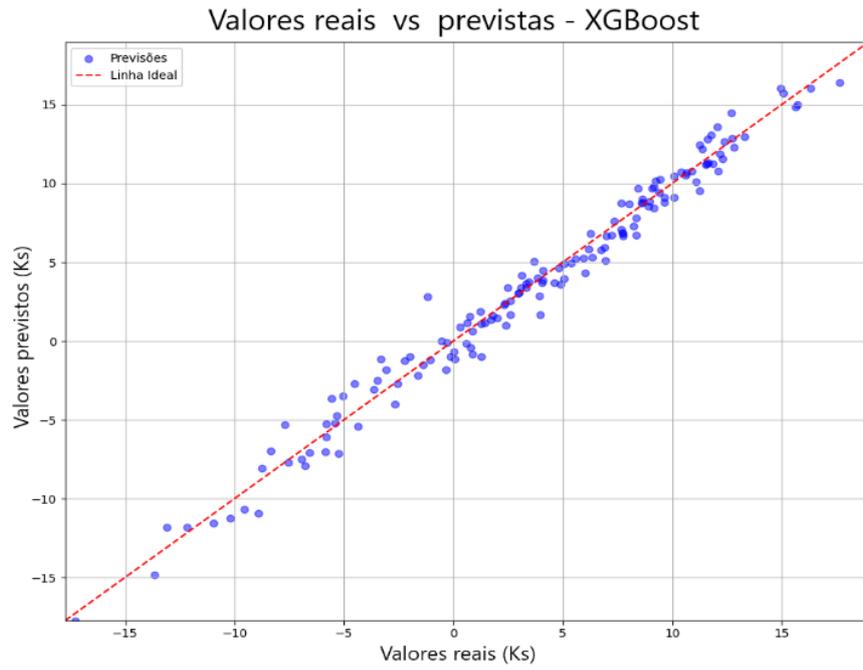
b)



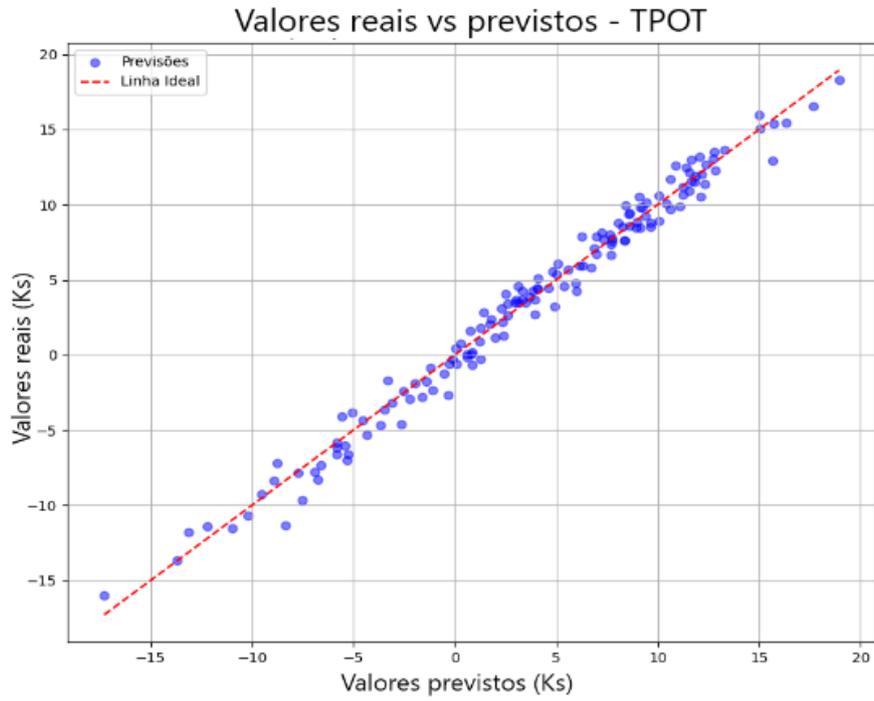
c)



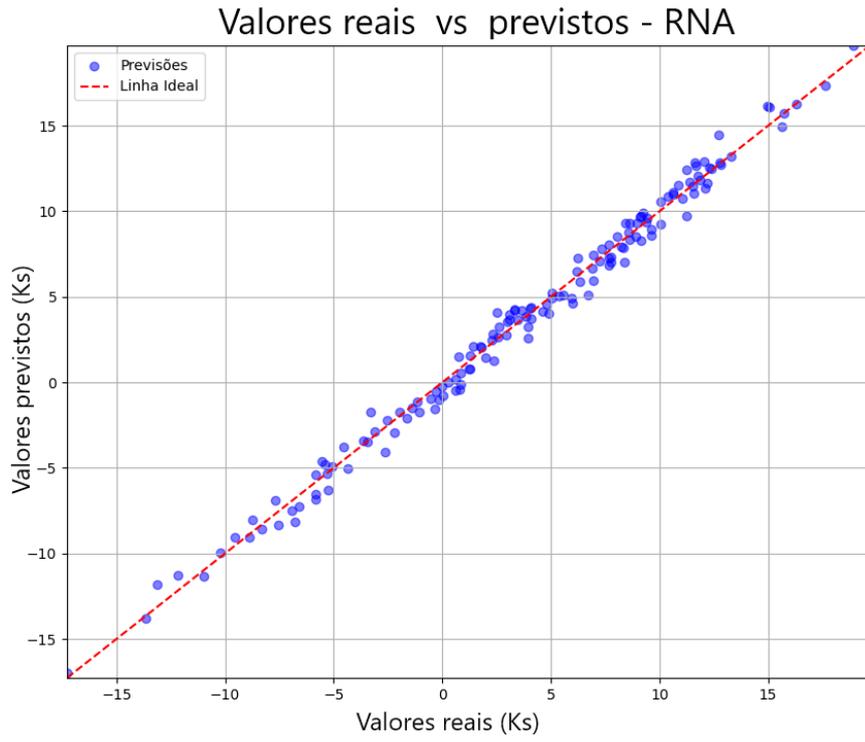
d)



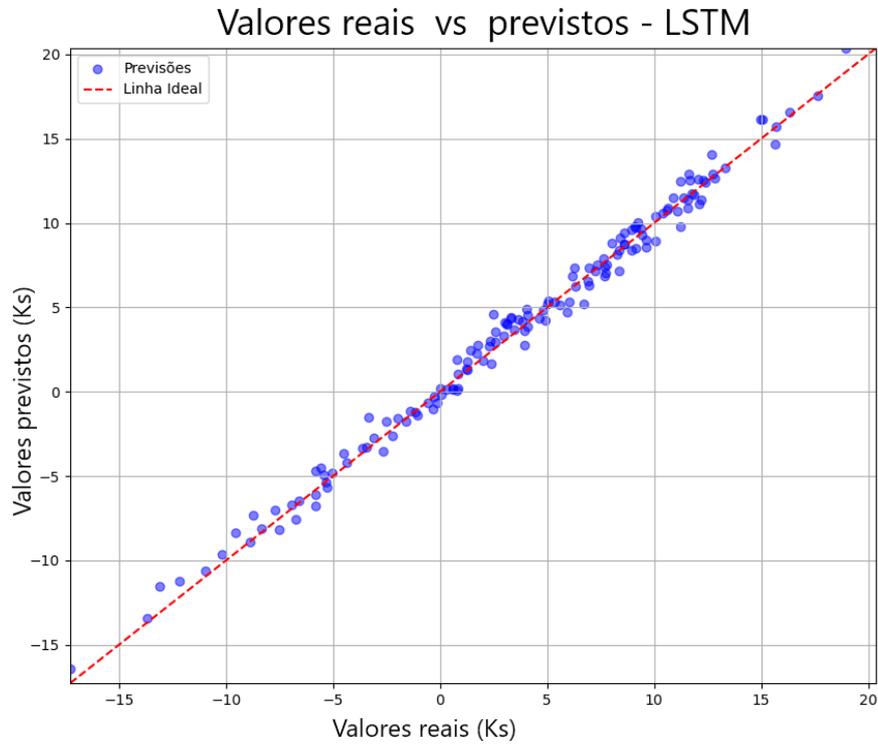
e)



f)



g)



Fonte: Autor (2024)

5.3 RESULTADOS E DISCUSSÃO

A perda de informação é inerente à representação de sistemas reais com alta variabilidade espacial e dinâmica multiescala, como em dados ambientais. Em contextos de amostragens pequenas, modelos de *machine learning* (ML) tendem a absorver incertezas estruturais sem explicitá-las (modelos “caixa-preta”), projetando-as como erros de generalização quando confrontados com padrões raros (extremos) ou mudanças de regime. Assim, a modelagem de sistemas complexos deve manter acurácia sem sacrificar a representatividade das relações entre variáveis — em particular caudas, assimetria e correlações que sustentam a física do processo.

Com esse objetivo, adotou-se um protocolo avaliativo em múltiplas camadas. Primeiro, procedeu-se à Análise Exploratória de Dados (AED) e aos testes de normalidade (*Shapiro–Wilk*, *Anderson–Darling*) e de duas amostras (*Kolmogorov–Smirnov*) para aferir fidelidade marginal (Tabelas 1 e 6; QQ-plots e ECDFs nas Figuras correspondentes). Em seguida, avaliou-se utilidade por transferência via TSTR/TRTR/TRTS, além da validação cruzada K-Fold dos modelos preditivos (Tabela 3). Por fim, realizou-se interpretação pós-hoc com SHAP para verificar se intervenções (p. ex., poda de outliers, transformações monotônicas, condicionamento por faixas) alteram a hierarquia de relevância entre preditores — um indício de quebra de suporte ou viés de seleção. Essa estratégia evita conclusões do tipo “melhorou a normalidade, logo o modelo é melhor”: normalidade univariada não implica utilidade preditiva; o que importa é o equilíbrio entre fidelidade e transferência.

A integração dessas camadas revelou um padrão consistente: (i) TabDDPM (difusão tabular) obteve a melhor utilidade em TSTR, com curva de aprendizado não-monótona (ver *Learning curve — TabDDPM*), mas ainda assim superior aos demais geradores em faixas relevantes de n ; (ii) GaussianCopula/CTGAN alcançou maior semelhança marginal (menores distâncias univariadas), porém com utilidade inferior; (iii) GAN/WGAN mostraram instabilidade (oscilação de perdas e indícios de colapso de modos), refletindo R^2 negativos/instáveis; (iv) VAE produziu amostras suavizadas e sub-dispersas, reduzindo amplitudes e caudas.

Com relação ao tratado na Tabela 1 (original) × Tabela 2 (sem poda; expandido por TabDDPM) observa-se compactação controlada e redução de assimetria/curtose típica da difusão:

- Ks: *skew* 5,86 → 0,97 e *kurt* 34,74 → 1,06; a cauda direita permanece sub-representada, mas os quantis centrais aderem melhor.

- Areia: $-1,24 \rightarrow -0,13$ (assimetria) e $3,01 \rightarrow -0,23$ (curtose) — distribuição mais equilibrada.
- Argila: $0,68 \rightarrow -0,42$ (assimetria) e $1,91 \rightarrow -0,95$ (curtose), com excesso de massa na faixa inferior nas figuras, demandando calibração por quantis e/ou condicionamento por faixas.
- Silte e Densidade do solo mantiveram comportamento próximo da simetria, com platocurtose discreta.

Essas mudanças explicam a não-monotonicidade da curva TSTR: ao inflar n sem correções dirigidas para Argila (cauda inferior) e Ks (cauda superior), o preditor treinado no sintético herda vieses e a transferência cai; com n maior, há recuperação por melhor cobertura do espaço.

Crucialmente, o conjunto com poda de outliers foi pior em todos os geradores: a retirada indiscriminada de extremos — sobretudo de Ks — quebrou o suporte do problema (os modelos deixam de “ver” regimes de alta condutividade), rebaixando a importância relativa desses preditores nos gráficos SHAP e piorando a generalização (K-Fold e TSTR). A recomendação é evitar poda e preferir estratégias robustas: transformações monotônicas por variável (p. ex., Yeo–Johnson/Box–Cox), condicionamento por faixas (notadamente Argila), reponderação de caudas em *mini-batches* (para Ks) e pós-ajuste por quantis (1–5% e 95–99%).

5.3.1 Análise do treinamento do método TabDDPM

Foi desenvolvida uma aplicação em Python® com *framework* Pytorch para criação dos modelos de aprendizado profundo, manipulação de dados e treinamento das redes neurais. Na definição do Gerador (*Generator*) foi utilizada uma rede neural sequencial, composta de camadas lineares, com funções de ativação LeakyReLU, com camadas de normalização BatchNorm1d. A função do gerador foi criar dados sintéticos com base em uma entrada de ruído aleatório.

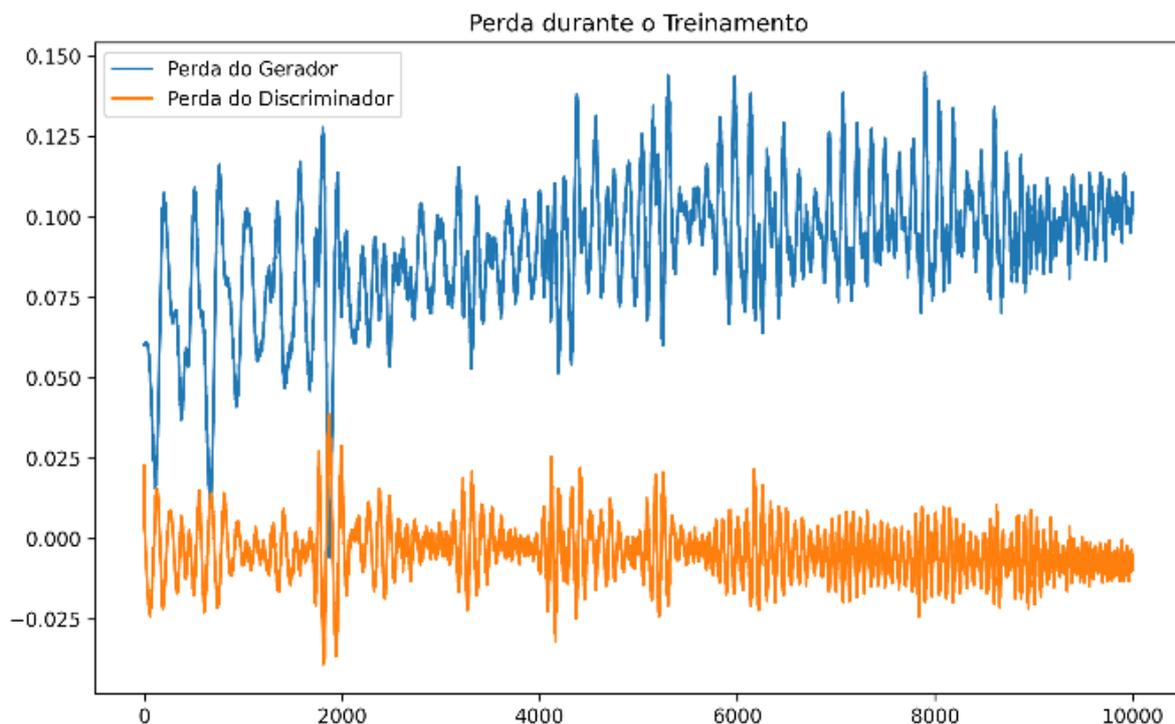
O discriminador (*Discriminator*) foi desenvolvido utilizando uma rede neural composta de camadas lineares e funções LeakyReLU e Sigmoid para saída, para distinguir entre dados reais e os gerados pelo gerador. Gerador e Discriminador foram instanciados para uma GPU, utilizando otimizadores Adam, configurados para ambos os modelos, com taxa de aprendizado e parâmetros de momento definidos. O treinamento ocorreu em um *loop* de várias épocas. O Discriminador foi treinado com uma amostra de dados reais, sendo gerado

um vetor de ruído e passado pelo gerador para criar dados sintéticos. A perda com função BCE (Binary Cross-Entropy Loss) foi então calculada.

Ao final, o gerador é treinado para melhorar sua capacidade de “enganar” o discriminador, minimizando a perda em que o discriminador classifica os dados gerados como reais. Em seguida, os dados são salvos em formatos específicos.

A escolha do gerador foi pautada por critérios hierarquizados: (i) *utilidade por transferência* para a tarefa preditiva (TSTR/TRTR/TRTS, com *learning curves* e IC95% via *bootstrap*), (ii) *estabilidade de treinamento* (histórico de perdas, ausência de colapso de modos), (iii) *fidelidade marginal e multivariada* (testes KS/AD/Shapiro–Wilk; QQ-plots/ECDF; KDE-2D), (iv) *conformidade físico-estatística* (preservação de suporte e plausibilidade de extremos) e (v) *simplicidade operacional* (robustez a poucos dados e poucos *tuning knobs*).

Figura 5 – Gráfico de geração de dados sintéticos



Fonte: Autor (2024)

Sintetizando os resultados comparativamente:

- TabDDPM (difusão tabular) — apresentou a melhor utilidade em TSTR, com *learning curve* não-monótona porém dominante em faixas relevantes de n (Figura “Learning curve — TabDDPM”). Mostrou boa aderência nos centrais de r_b , Areia, Silte, e vieses localizados: encolhimento de cauda direita em K_s e excesso de massa na cauda

inferior de Argila. Tais vieses explicam a queda intermediária da curva e a recuperação em n maiores; são mitigáveis com transformações monotônicas, condicionamento por faixas (Argila) e quantile matching nas bandas 1–5% e 95–99%. O treinamento foi estável (perdas suaves) e a amostragem, reproduzível.

- GaussianCopula/CTGAN (família SDV) — obtiveram alta fidelidade univariada (menores distâncias marginais, p-values mais favoráveis), porém utilidade inferior na tarefa, sobretudo quando o alvo demanda caudas (Ks).
- GAN/WGAN — históricos de perda indicaram discriminador dominante e risco de colapso de modos, coerentes com R^2 negativos/instáveis em TSTR. Mesmo com ajustes (TTUR, *label smoothing*, *instance noise*), a amostra pequena e a dependência tabular penalizaram a estabilidade.
- VAE — gerou amostras suavizadas e sub-dispersas (amplitudes encurtadas), o que reduziu a transferência quando o regressor requer extremos.

5.3.2 Análise do conjunto de dados originais e expandidos

A implementação foi realizada em Python/PyTorch, modelando o TabDDPM como um denoiser MLP que estima o ruído (ou *score*) ao longo de uma trajetória de difusão com T passos e β -*schedule* suave. O pré-processamento incluiu padronização e, quando indicado, transformações monotônicas (e.g., Yeo–Johnson) — especialmente em Ks — com inversão após a geração. O treinamento utilizou Adam/AdamW, gradient clipping e EMA dos pesos; a amostragem foi por *ancestral sampling* (com opção DDIM para *fast sampling*). Essa configuração, consagrada em difusão tabular, maximiza estabilidade e cobertura do espaço latente.

A curva de aprendizado TSTR do TabDDPM mostra pico em ~ 200 amostras, queda em torno de 1000 e recuperação em 2000. As figuras univariadas (histogramas, QQ-plots e scatters) mostram boa aderência em rb, Areia e Silte, sub-amostragem de caudas altas em Ks e acúmulo na cauda inferior de Argila — exatamente os efeitos que se propagam ao preditor e explicam a forma da curva. Três ajustes práticos elevam a robustez do difusor neste caso: (i) Condicionamento por faixas de Argila (baixa/média/alta), para mitigar mistura de modos; (ii) Reponderação de cauda (oversampling dos quantis $>95\%$) no *score matching* de Ks, para restaurar extremos; (iii) Pós-ajuste por *quantile mapping* nas bandas 1–5% e 95–99%, preservando ordem e correlações.

Para comparação histórica, a Figura 5 (perdas do GAN) mostra discriminador dominante: sua perda cai rapidamente e se mantém baixa, enquanto a perda do gerador sobe e oscila em patamares altos — padrão compatível com colapso de modos e gradientes pobres, o que explica os R^2 negativos/instáveis do GAN em TSTR. Esse contraste reforça a escolha pela difusão tabular neste estudo.

A decisão e especificação operacional considerou os critérios de seleção (utilidade por transferência, estabilidade, fidelidade uni/multivariada, plausibilidade físico-estatística e simplicidade operacional), o método de expansão adotado é o TabDDPM. A decisão se ancora em: (1) desempenho superior em TSTR ao longo da *learning curve*; (2) estabilidade de treinamento sem colapso; (3) capacidade de correção dirigida por variável, permitindo alinhar caudas e misturas sem sacrificar correlações. Para reprodutibilidade: denoiser MLP em PyTorch, Adam/AdamW, gradient clipping, EMA, β -schedule suave, *sampling* ancestral (ou DDIM), tamanhos sintéticos {50, 100, 200, 500, 1000, 2000} e avaliação por TSTR/TRTR/TRTS, K-Fold, KS/AD/SW, QQ/ECDF/KDE-2D e SHAP.

A análise do conjunto de dados originais e expandidos mostrou que os conjuntos originais com pontos extremos (44 amostras) e com poda (35 amostras) foram avaliados visando identificar inconsistências e distorções (medição, representação, completude), que afetam a qualidade dos dados e a acurácia de modelos de ML. Em amostras pequenas, a poda prévia pode eliminar informação sobre distribuição, espacialidade e dependências entre variáveis, prejudicando o desempenho de algoritmos — motivo pelo qual sua adoção deve ser investigada previamente.

A inspeção por IQR evidenciou outliers pontuais, mas a maioria das amostras concentra-se em intervalos consistentes. Em Areia, há poucos outliers inferiores; em Ks, há extremos superiores condizentes com micro-heterogeneidades do meio poroso; em Argila, observam-se valores baixos recorrentes. Esses padrões são naturais em variáveis físico-químicas do solo e, por si sós, não justificam poda ampla.

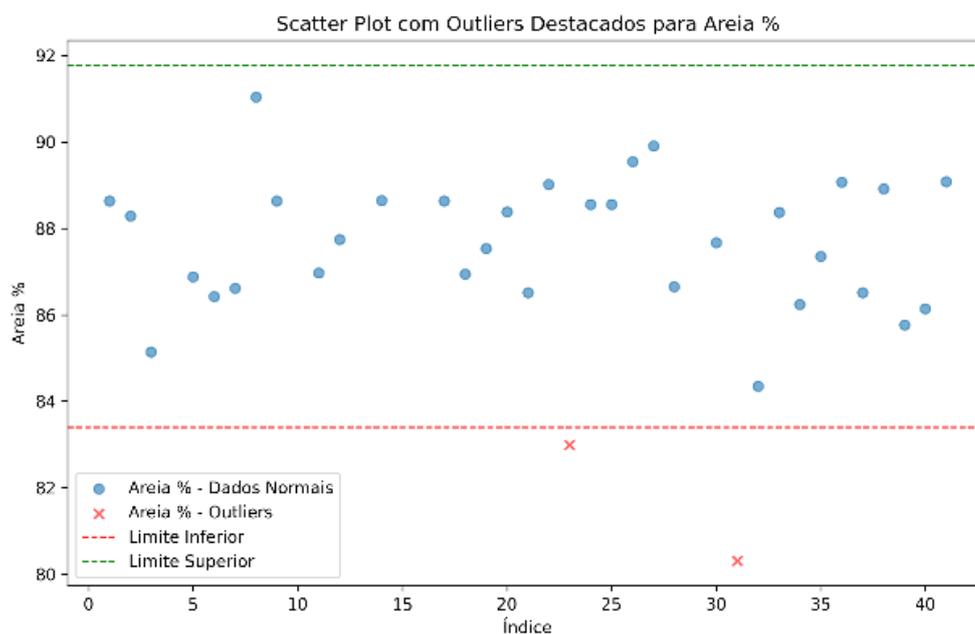
O confronto Tabela 1 (original) × Tabela 2 (sem poda; TabDDPM) indica:

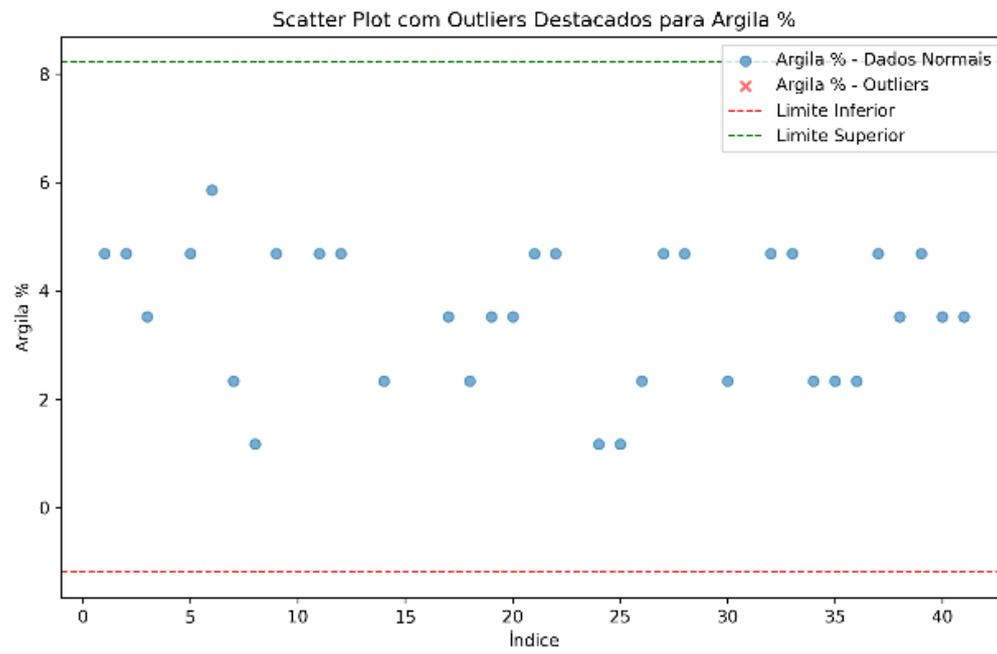
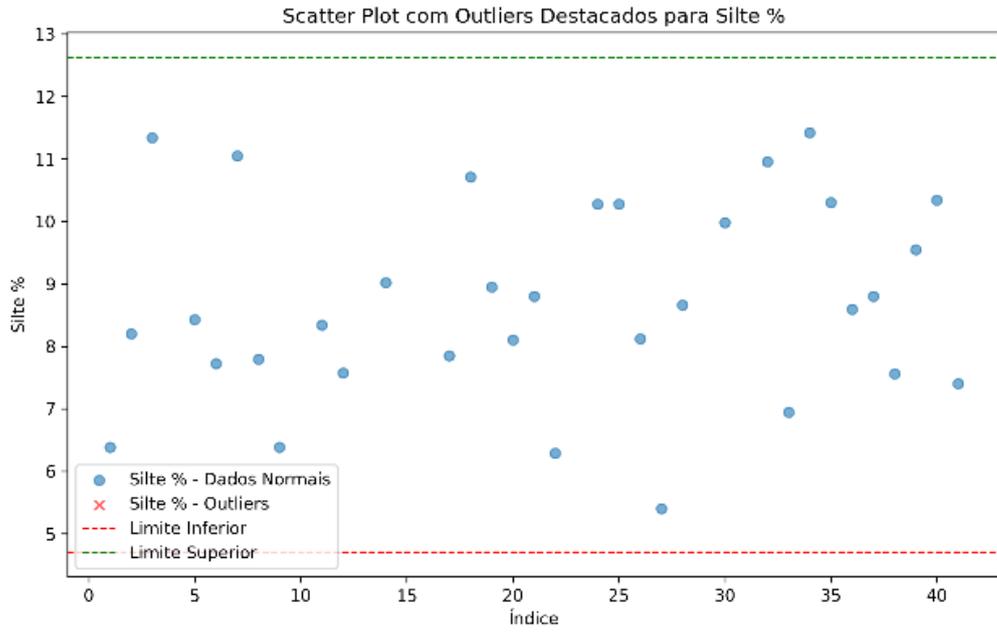
- Areia: média estável (87,50% → 87,71%), assimetria/curtose reduzidas (-1,24 → -0,13; 3,01 → -0,23), sugerindo distribuição mais equilibrada; a menor dispersão tende a facilitar algoritmos sensíveis a assimetrias (e.g., RNA) sem comprometer correlações com Ks.

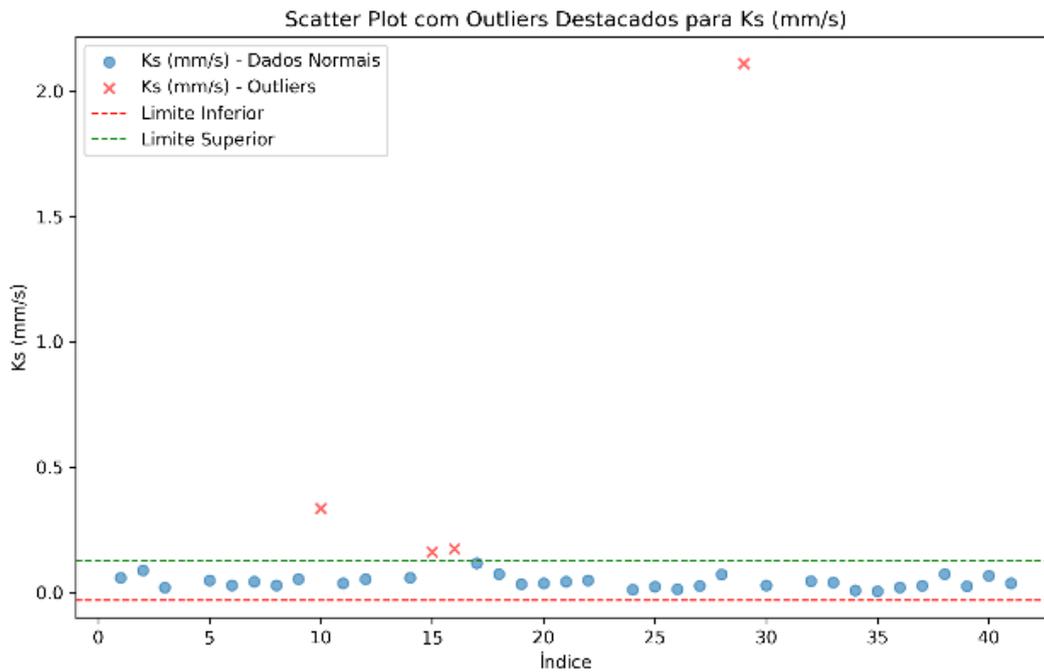
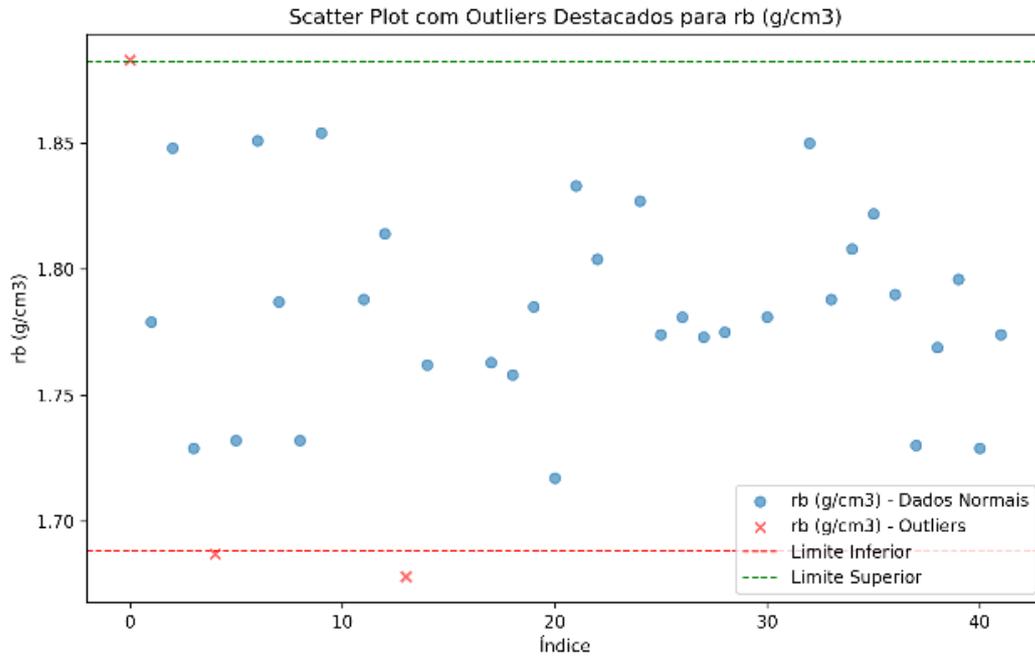
- Silte: média praticamente inalterada (8,72% → 8,71%), platocurtose leve (-0,75 → -0,78), distribuição estável e pouco sensível a outliers.
- Argila: média levemente menor (3,80% → 3,58%), amplitude reduzida (8,21% → 4,69%), assimetria/curtose diminuídas (0,68 → -0,42; 1,91 → -0,95); todavia, as figuras revelam acúmulo na cauda inferior, sugerindo condicionamento por faixas e quantile matching para preservar regimes de maior argilosidade.
- Densidade (rb): consistência alta (média ~1,78), leve achatamento (-0,24 → -0,68), preservando estrutura.
- Ks: média menor ($\approx 0,11 \rightarrow 0,04$), forte redução de assimetria/curtose (5,86 → 0,97; 34,74 → 1,06); a cauda direita continua sub-representada, motivo para aplicar reponderação de cauda e eventual tail repair por GPD nos quantis superiores.

Os resultados obtidos destacaram presença de *outliers* nas variáveis preditoras do solo de estudo (Figura 6) podem influir negativamente na acurácia de modelagem IA, mas há necessidade de se investigar se sua poda não acarretará perda de representatividade dos dados, buscando-se o equilíbrio entre distribuição e preservação de variabilidade natural das amostras.

Figura 6 – Visualização das variáveis de estudo nos limites superior e inferior (*scatter plots*), caracterizando pontos extremos ou *outliers* do conjunto amostral experimental original.







Fonte: Autor (2024)

Na avaliação comparativa dos conjuntos expandidos por múltiplas técnicas (GAN, WGAN, VAE, SDV/CTGAN/GaussianCopula e TabDDPM), os resultados indicaram predominância do TabDDPM no critério de utilidade por transferência (TSTR/TRTR/TRTS)

e estabilidade de treinamento, com *learning curve* superior em faixas relevantes de amostras sintéticas. Os métodos GAN/WGAN exibiram instabilidades e sinais de colapso de modos, o VAE produziu amostras sub-dispersas (amplitudes encurtadas), e os modelos SDV/CTGAN/GaussianCopula obtiveram boa fidelidade marginal mas menor utilidade na tarefa preditiva. Em razão desse conjunto de evidências, o método de expansão adotado neste estudo é o TabDDPM (difusão tabular), utilizado sem poda de *outliers*.

Tabela 4 – Análise descritiva do conjunto de dados sintéticos sem poda, expandido por TabDDPM.

	Variável	Média	Desvio-padrão	Mediana	Mínimo	Máximo	Amplitude
Originais	Areia (%)	87,8511	1,5305	88,2900	84,3500	91,0400	6,6900
	Silte (%)	8,6483	1,5601	8,4300	5,4000	11,4200	6,0200
	Argila (%)	3,5166	1,3046	3,5200	1,1700	5,8600	4,6900
	rb (g/cm ³)	1,7849	0,0440	1,7810	1,6870	1,8830	0,1960
	Ks (mm/s)	0,0422	0,0244	0,0373	0,0049	0,1170	0,1121
Sintéticos	Areia (%)	87,7507	1,2184	87,6500	84,3500	91,0400	6,6900
	Silte (%)	8,5961	1,2097	8,5900	5,4000	11,4200	6,0200
	Argila (%)	3,8830	1,1396	4,1800	1,1700	5,7900	4,6200
	rb (g/cm ³)	1,7816	0,0364	1,7820	1,6870	1,8830	0,1960
	Ks (mm/s)	0,0346	0,0124	0,0349	0,0049	0,0727	0,0678
Variável	Δ Média	Δ Desvio-padrão	Δ Mediana	Δ Mínimo	Δ Máximo	Δ Amplitude	Encolhimento de amplitude
Areia (%)	-0,1005	-0,3121	-0,6400	0,0000	0,0000	0,0000	0,0000
Silte (%)	-0,0522	-0,3504	0,1600	0,0000	0,0000	0,0000	0,0000
Argila (%)	0,3665	-0,1650	0,6600	0,0000	-0,0700	-0,0700	1,4925
rb (g/cm ³)	-0,0033	-0,0076	0,0010	0,0000	0,0000	0,0000	0,0000
Ks (mm/s)	-0,0076	-0,0120	-0,0024	0,0000	-0,0443	-0,0443	39,5347

Fonte: Autor (2024)

A leitura dos QQ-plots (Figura 7) confirma essas conclusões. Areia e Silte mostram boa aderência à reta de referência, com pequenas curvaturas nas extremidades; rb mantém comportamento próximo da normalidade; Argila exibe concavidade em quantis baixos; e Ks alinha-se melhor ao centro, mas conserva desvios na cauda superior. Essa evidência visual é consistente com as estatísticas descritivas: quando se compara a Tabela 1 (original) com a Tabela 2 (sem poda; TabDDPM) observa-se redução pronunciada de assimetria/curtose — por exemplo, em Ks (*skew* 5,86 → 0,97; *kurt* 34,74 → 1,06) e em Areia (-1,24 → -0,13; 3,01 → -0,23) — junto de acúmulo na faixa inferior de Argila (0,68 → -0,42; 1,91 → -0,95). Esses padrões explicam a não-monotonicidade da *learning curve* TSTR: se aumentamos o número de amostras sintéticas n sem correções dirigidas para Argila (cauda inferior) e Ks (cauda superior), a transferência degrada por viés de cauda; com n maior e/ou calibração por quantis, a cobertura do espaço melhora e a curva se recupera.

Tabela 5 – Testes de normalidade para o conjunto original (sem poda) de *outliers*.

Variável	Kolmogorov–Smirnov (p)	Anderson–Darling (estat.)	Shapiro–Wilk (p)	Conjunto
Areia (%)	0,5797	0,4275	0,5768	Original (baseline)
Silte (%)	0,7563	0,4296	0,3644	Original (baseline)
Argila (%)	0,0252	1,9841	0,0007	Original (baseline)
rb (g/cm ³)	0,7360	0,3794	0,6926	Original (baseline)
Ks (mm/s)	0,7755	0,5044	0,0774	Original (baseline)

Fonte: Autor (2024)

Os testes de aderência (Tabelas 5 e 6) mostraram, na maior parte das variáveis, estatísticas reduzidas e p-valores não significativos no Kolmogorov–Smirnov (KS), corroboradas por Anderson–Darling (AD) e Shapiro–Wilk (SW). Esse padrão é consistente com as inspeções visuais (histogramas, QQ-plots e ECDFs): Areia, Silte e rb apresentam bom alinhamento aos quantis teóricos nos centros, com desvios discretos de cauda; Argila conserva excesso de massa na cauda inferior e Ks mantém sub-representação da cauda direita, apesar de melhora substancial em relação ao conjunto original. Em outras palavras, a difusão compacta caudas extremas e regulariza a forma marginal sem descaracterizar as relações principais.

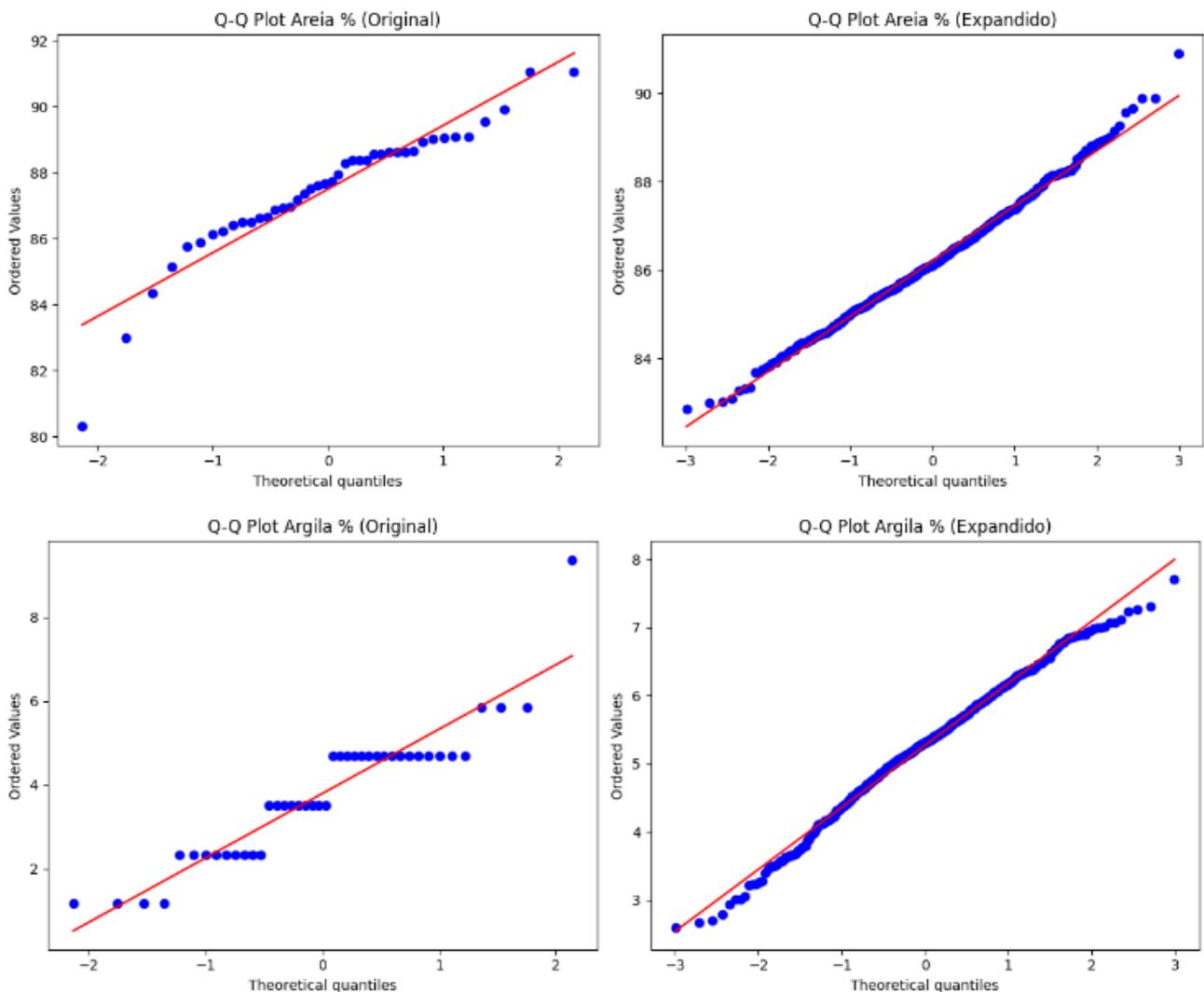
Tabela 6 – Testes de normalidade para o conjunto sintético TabDDPM (sem poda) de *outliers*.

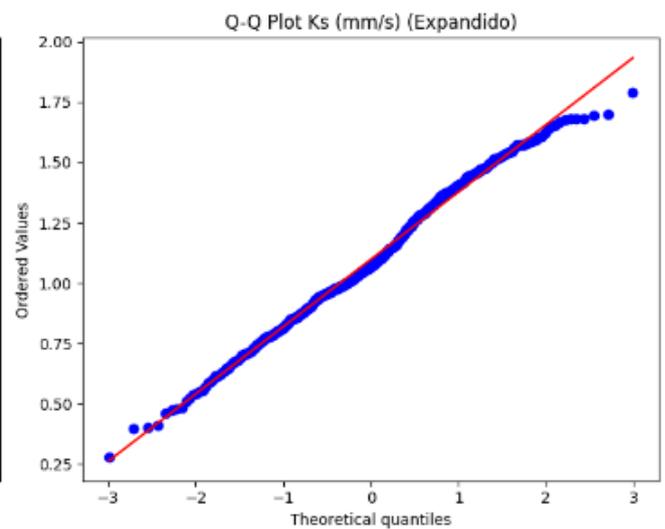
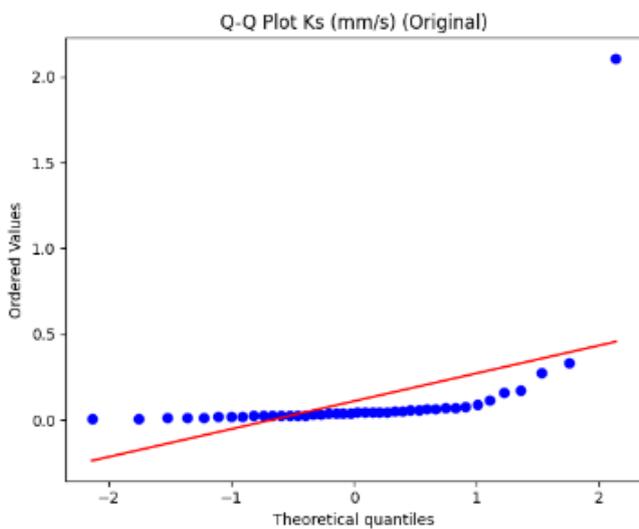
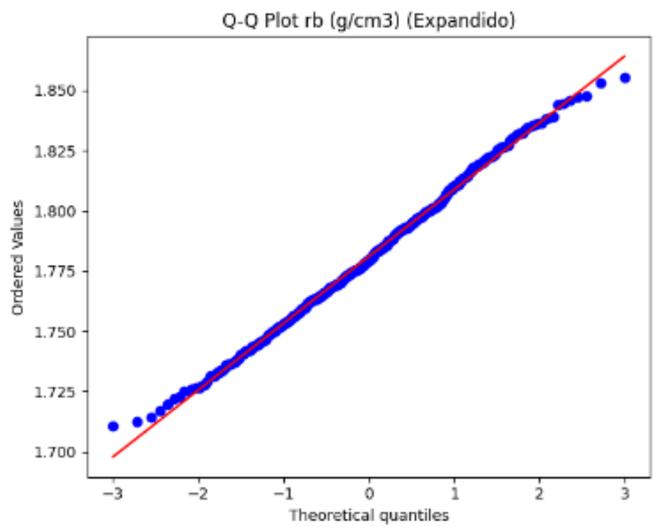
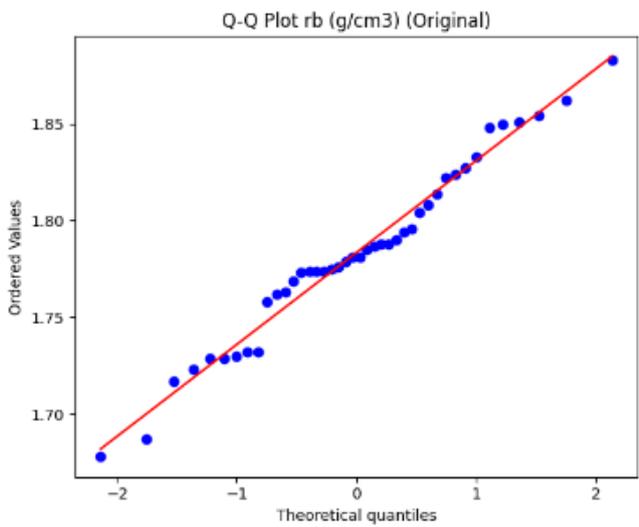
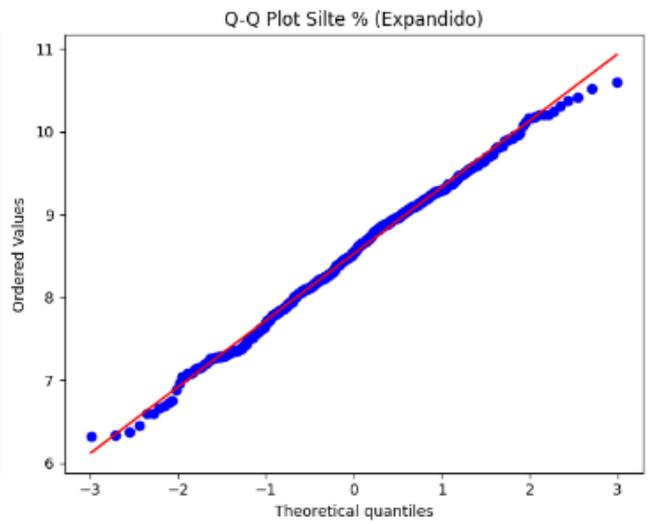
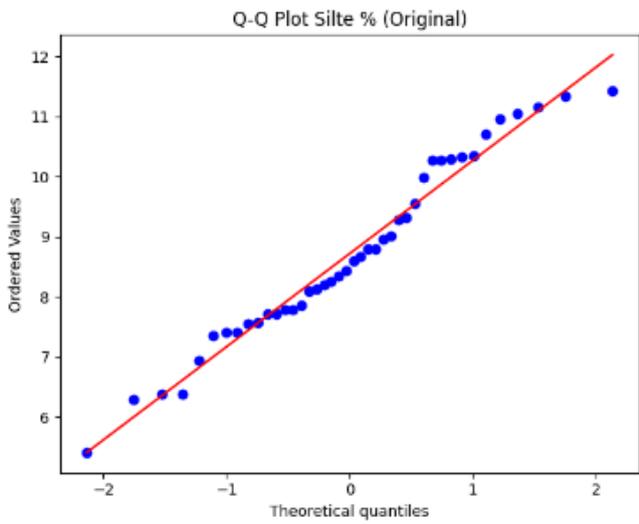
Variável	Kolmogorov–Smirnov (p)	Anderson–Darling (estat.)	Shapiro–Wilk (p)	Conjunto
Areia (%)	0,6237	0,6772	0,0001	Sintético (SDV_GaussianCopula)
Silte (%)	0,0000	16,4189	0,0000	Sintético (SDV_GaussianCopula)
Argila (%)	0,0000	23,1756	0,0000	Sintético (SDV_GaussianCopula)
rb (g/cm ³)	0,4482	1,0313	0,0001	Sintético (SDV_GaussianCopula)
Ks (mm/s)	0,0000	12,6979	0,0000	Sintético (SDV_GaussianCopula)
Areia (%)	0,0000	27,5553	0,0000	Sintético (VAE)
Silte (%)	0,2184	1,6282	0,0010	Sintético (VAE)
Argila (%)	0,1097	1,2610	0,0014	Sintético (VAE)
rb (g/cm ³)	0,2284	0,8692	0,2390	Sintético (VAE)
Ks (mm/s)	0,6857	1,0227	0,0024	Sintético (VAE)
Areia (%)	0,0499	2,9448	0,0000	Sintético (GAN)
Silte (%)	0,1065	2,3267	0,0000	Sintético (GAN)
Argila (%)	0,0411	3,0332	0,0000	Sintético (GAN)
rb (g/cm ³)	0,0323	2,8629	0,0000	Sintético (GAN)
Ks (mm/s)	0,2791	1,1246	0,0001	Sintético (GAN)
Areia (%)	0,0858	1,4404	0,0026	Sintético (TabDDPM)
Silte (%)	0,6748	0,7054	0,0017	Sintético (TabDDPM)
Argila (%)	0,0000	30,0869	0,0000	Sintético (TabDDPM)
rb (g/cm ³)	0,8089	0,1961	0,2893	Sintético (TabDDPM)
Ks (mm/s)	0,5795	0,5949	0,0218	Sintético (TabDDPM)

Fonte: Autor (2024)

Outro aspecto relevante é a presença de outliers no conjunto original. Em amostragens pequenas, a poda desses pontos quebra o suporte do problema — notadamente para Ks, cujo regime de alta condutividade carrega informação crítica. Isso se refletiu em p-valores marginais para algumas variáveis e pior desempenho em K-Fold/TSTR quando a poda foi aplicada, além de reordenar as contribuições SHAP (menor importância de preditores associados a regimes extremos). Por essa razão, não utilizamos poda no fluxo final; optamos por estratégias robustas e direcionadas: transformações monotônicas (Yeo–Johnson/Box–Cox), condicionamento por faixas (especialmente para Argila), reponderação de caudas em *mini-batches* (Ks) e pós-ajuste por quantile mapping nas bandas 1–5% e 95–99%. Tais medidas preservam extremos plausíveis, mantêm correlações e melhoram a transferência.

Figura 7 – Visualização gráfica QQ-plots das variáveis de estudo nos conjuntos de dados originais e expandidos.





A avaliação final da expansão do conjunto amostral confirma que o tamanho reduzido da amostra original limita a replicação fiel das relações inter-variáveis e da heterogeneidade espacial do sistema, sobretudo quando se aplica poda de outliers antes de gerar sintéticos. Embora a poda suavize caudas extremas, ela remove regimes informativos (raros, porém plausíveis), o que altera a estrutura de dependências e quebra o suporte dos dados (especialmente de K_s), degradando a capacidade de generalização dos modelos (TSTR/K-Fold) e reordenando a importância relativa dos preditores (SHAP). Por isso, neste estudo, adotamos a expansão sem poda e selecionamos o TabDDPM como método gerador principal, por combinar estabilidade de treinamento, fidelidade estatística e utilidade por transferência superiores às alternativas testadas (GAN, WGAN, VAE, SDV/CTGAN/GaussianCopula).

As Figuras de boxplot (Areia, Silte, Argila, pb e K_s) e os scatters vs K_s reforçam esses pontos.

(i) Em Areia e Silte, observa-se variabilidade moderada com dispersão predominantemente central e poucos valores extremos; os boxplots padronizados (0–1) mostram quartis equilibrados, o que coaduna com os QQ-plots (boas aderências no centro e pequenas curvaturas nas caudas). (ii) Em Argila, há acúmulo na cauda inferior — visível no QQ-plot e no boxplot — indicando que o difusor tende a sub-amostrar faixas altas quando não condicionado; é um ponto de atenção para calibração dirigida (condicionamento por faixas; *quantile mapping* no 1–5%). (iii) Em pb, a distribuição é estreita e estável, com poucos extremos; isso favorece modelos que dependem de métricas de distância ou margens (SVM/kNN), desde que a escala seja harmonizada. (iv) Em K_s , os boxplots evidenciam sub-representação da cauda direita no sintético — coerente com o decréscimo observado em *skew/kurt* (Tabela 2) e com os desvios dos QQ-plots nas ordens superiores. Esse comportamento é esperado em difusão tabular quando caudas são raras no conjunto original; a correção prática envolve reponderação de cauda nos *mini-batches* de *score matching* e, se necessário, ajuste por cauda pesada (e.g., *tail repair* com GPD) restrito aos quantis >95%.

Os scatters multivariados vs K_s (painéis “(a) com outliers” e “(b) sem outliers/padronizado”) são particularmente elucidativos: com poda, nota-se estreitamento artificial das nuvens e perda de inclinações nas relações, ao passo que a versão sem poda e expandida com TabDDPM preserva gradientes e faixas de resposta mais amplas, ainda que com menor densidade nos extremos. Esse equilíbrio — *regularização da forma*

marginal sem apagar estruturas — é precisamente o que favoreceu o desempenho TSTR do TabDDPM nas faixas de n onde sua *learning curve* foi dominante.

Quanto ao impacto na modelagem de IA, os resultados são coerentes com a literatura e com nossas métricas:

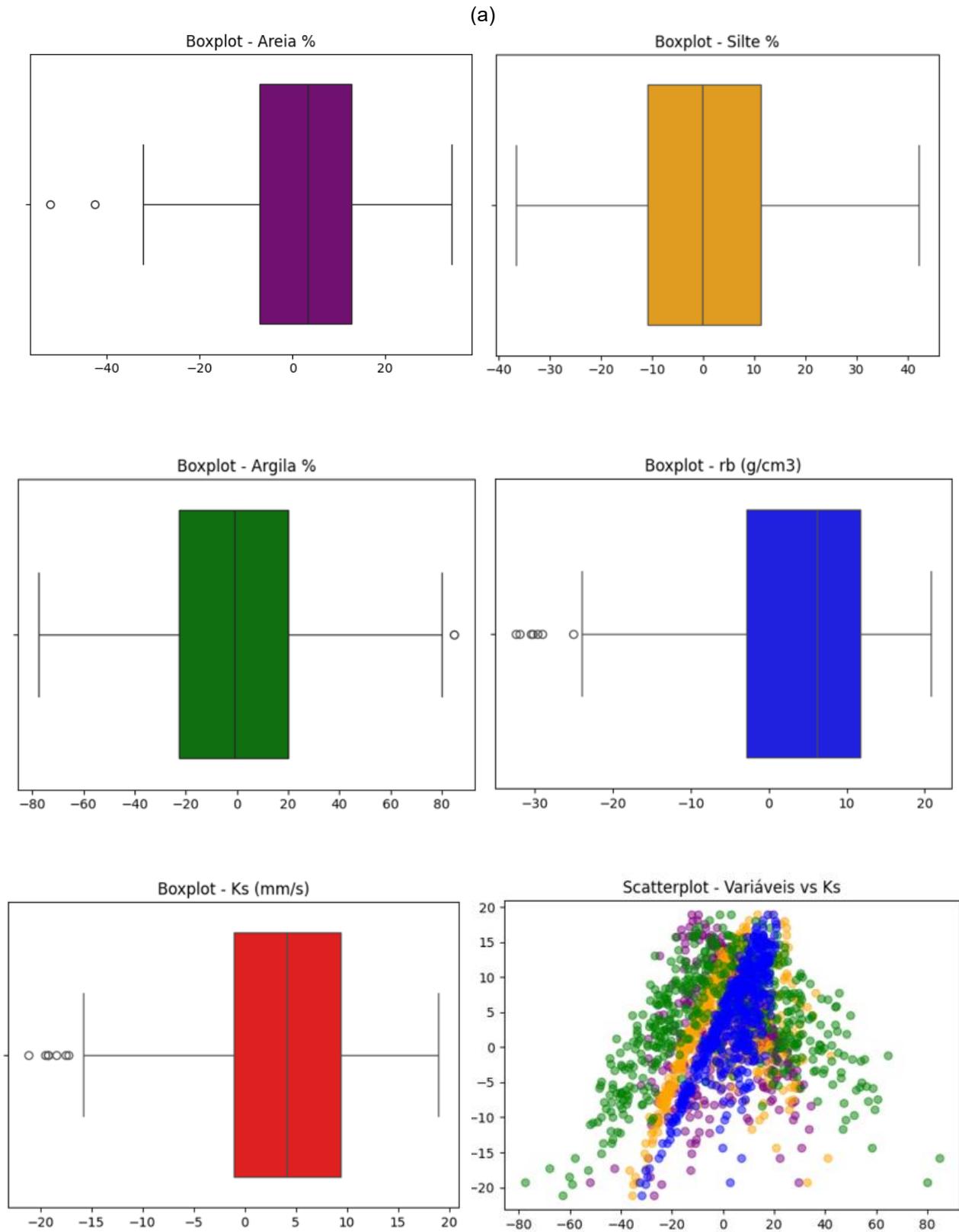
- Modelos em árvore (Random Forest, XGBoost) mostraram-se robustos à heterogeneidade residual, beneficiando-se da melhor cobertura do espaço sintético;
- Redes neurais (MLP/LSTM) reagiram positivamente à redução de assimetria/curtose em Areia/Silte/pb, desde que mantidos os regimes de cauda que “ensinam” o mapeamento para Ks;
- SVM/kNN foram sensíveis à sub-representação de extremos (Argila baixa, Ks alto), reforçando a necessidade de condicionamento (Argila) e reponderação (Ks) ao treinar o difusor.

Dessa análise decorre um protocolo de boas práticas para este domínio:

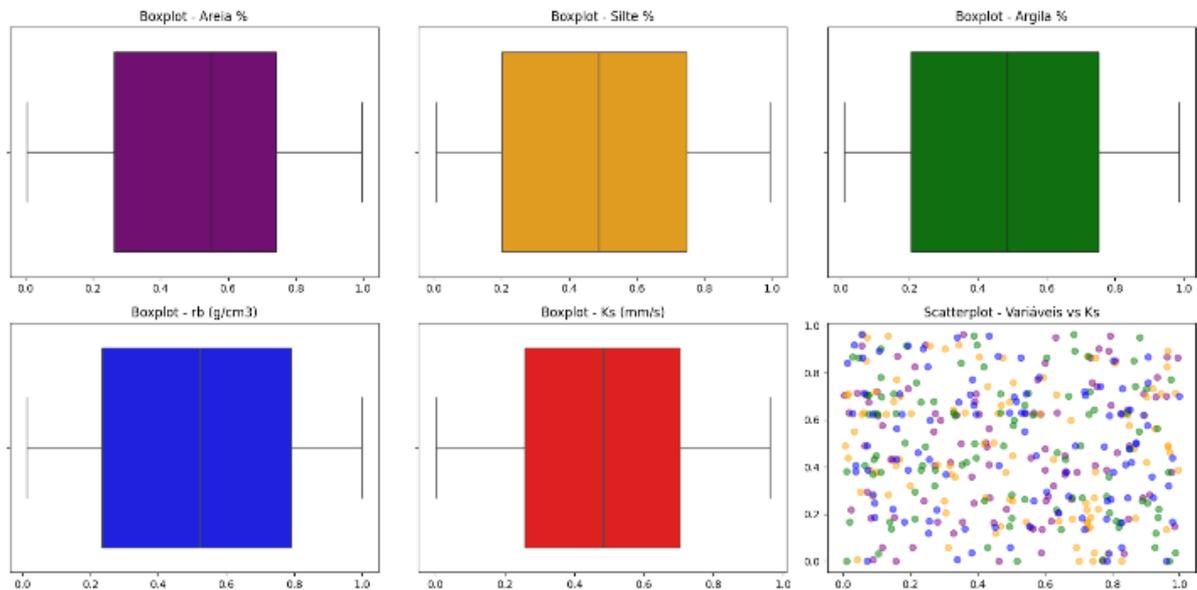
1. Não podar outliers a priori;
2. Transformações monotônicas por variável (e.g., Yeo–Johnson para Ks/Argila), com inversão após a geração;
3. Condicionamento do difusor por faixas de Argila (baixa/média/alta) para evitar mistura de modos;
4. Reponderação de cauda de Ks durante o treinamento do TabDDPM;
5. Pós-calibração por quantis (1–5% e 95–99%) para alinhamento fino sem distorcer correlações;
6. Validação por TSTR/TRTR/TRTS com *bootstrap* de IC95% e verificação de consistência SHAP.

Em síntese, a pequena amostra condiciona o problema, mas o TabDDPM sem poda mostrou-se a estratégia mais equilibrada para gerar sintéticos úteis do ponto de vista estatístico e preditivo. As distorções residuais — Argila em cauda inferior e Ks em cauda superior — são localizadas e mitigáveis com o *pipeline* acima, sem sacrificar a representatividade física do sistema.

Figura 8 – Visualização das variáveis de estudo nos limites superior e inferior por *scatter plots*, caracterizando pontos extremos (outliers) e distribuição central: (a) conjunto com poda de outliers (efeito de estreitamento e perda de suporte); (b) conjunto sem poda, expandido por TabDDPM (preservação de gradientes e faixas de resposta, com leve sub-representação de caudas).



(b)



Fonte: Autor (2024)

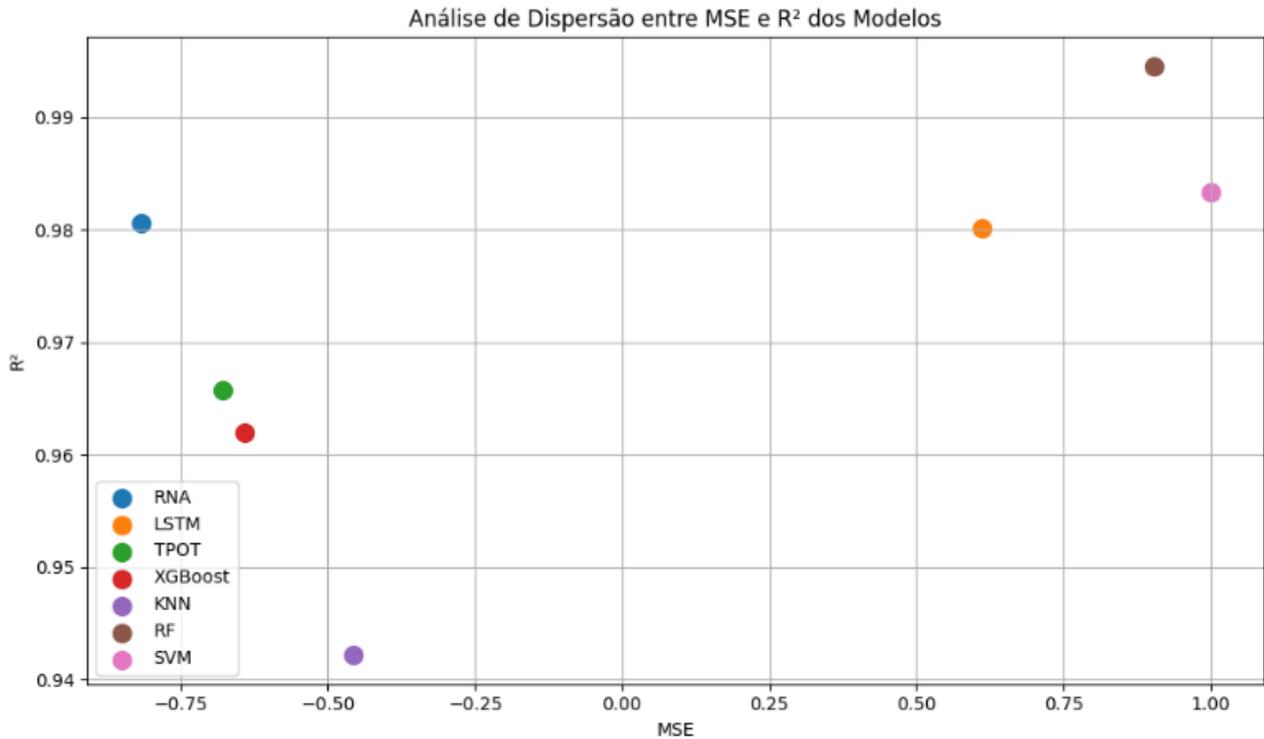
5.3.3 Análises dos resultados da predição com Redes Neurais Artificiais

O algoritmo de RNA apresentou melhor acurácia dentre todos os modelos analisados, considerando como métricas de MSE, RMSE, MAE e R^2 (Tabela 3). Os resultados obtidos confirmam os cenários previstos nas análises dos dados originais e expandidos sem poda de *outliers* por GAN, com alta complexidade relacional, heterogêneos, com vantagens na modelagem por alguns modelos. A padronização de procedimentos de desenvolvimento dos modelos, como definição automática de hiperparâmetros, sem refinamentos caso a caso, sem ajustes sucessivos dos dados expandidos, mas garantindo ocorrência de erros fatais, overfitting ou outras inconsistências.

Uma análise geral dos resultados (Figura 9) mostra que os algoritmos neurais conseguiram melhor estimar os valores de condutividade hidráulica saturada, frente aos dados disponibilizados, mais balanceados. Uma análise de dispersão entre o erro quadrático médio (MSE) e a explicabilidade de performance (R^2) dos modelos foi elaborada, avaliando correção e consistência. RNA e LSTM estão no quadrante ideal, com baixo valor de MSE e alto R^2 . SVM e RF não conseguiram ótimas estimativas, sugerindo necessidades de ajustes nos parâmetros e/ou dificuldade em pontos com variabilidade extrema, presentes nos dados expandidos. O algoritmo de KNN apresentou menor desempenho, apresentando

MSE alto e R^2 abaixo da maioria, indicando baixa adaptabilidade aos dados apresentados, além de necessidades de ajustes finos em seus parâmetros.

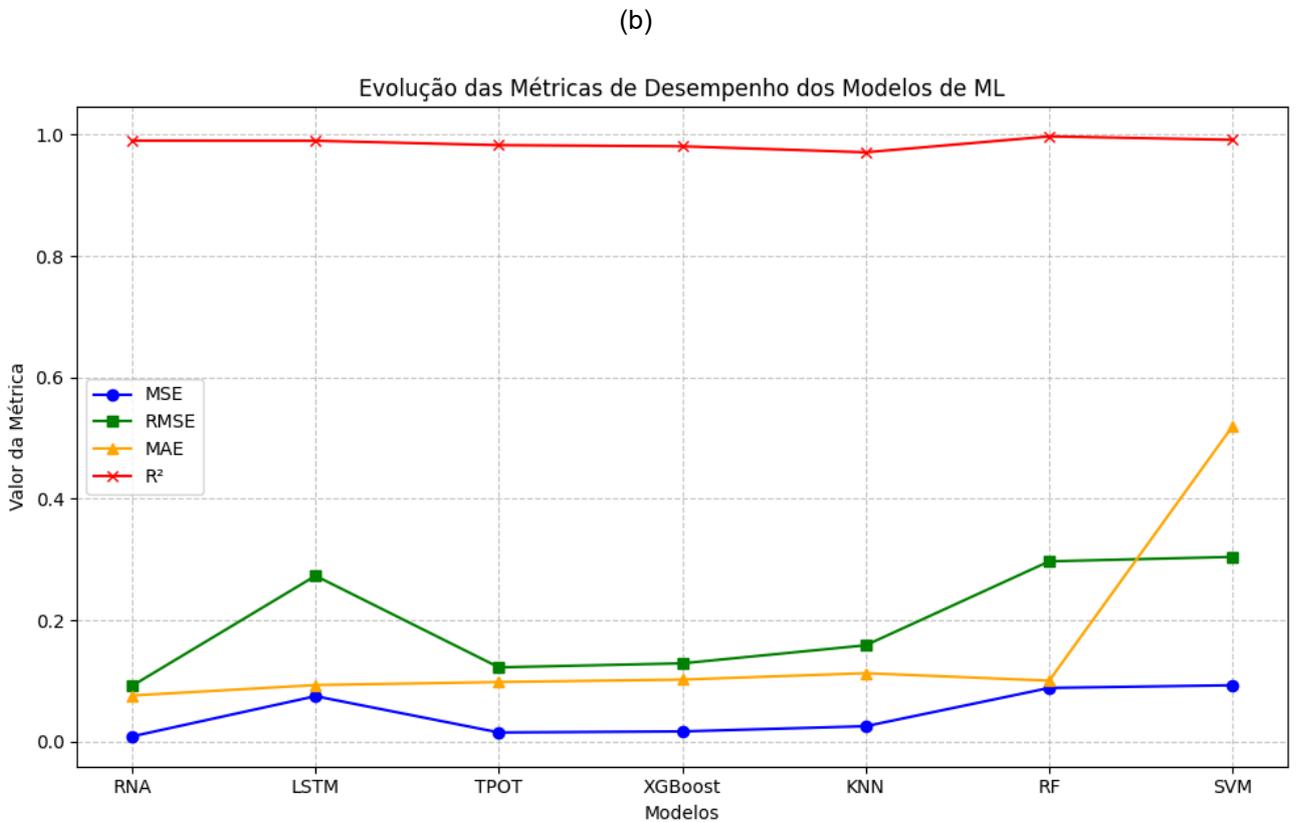
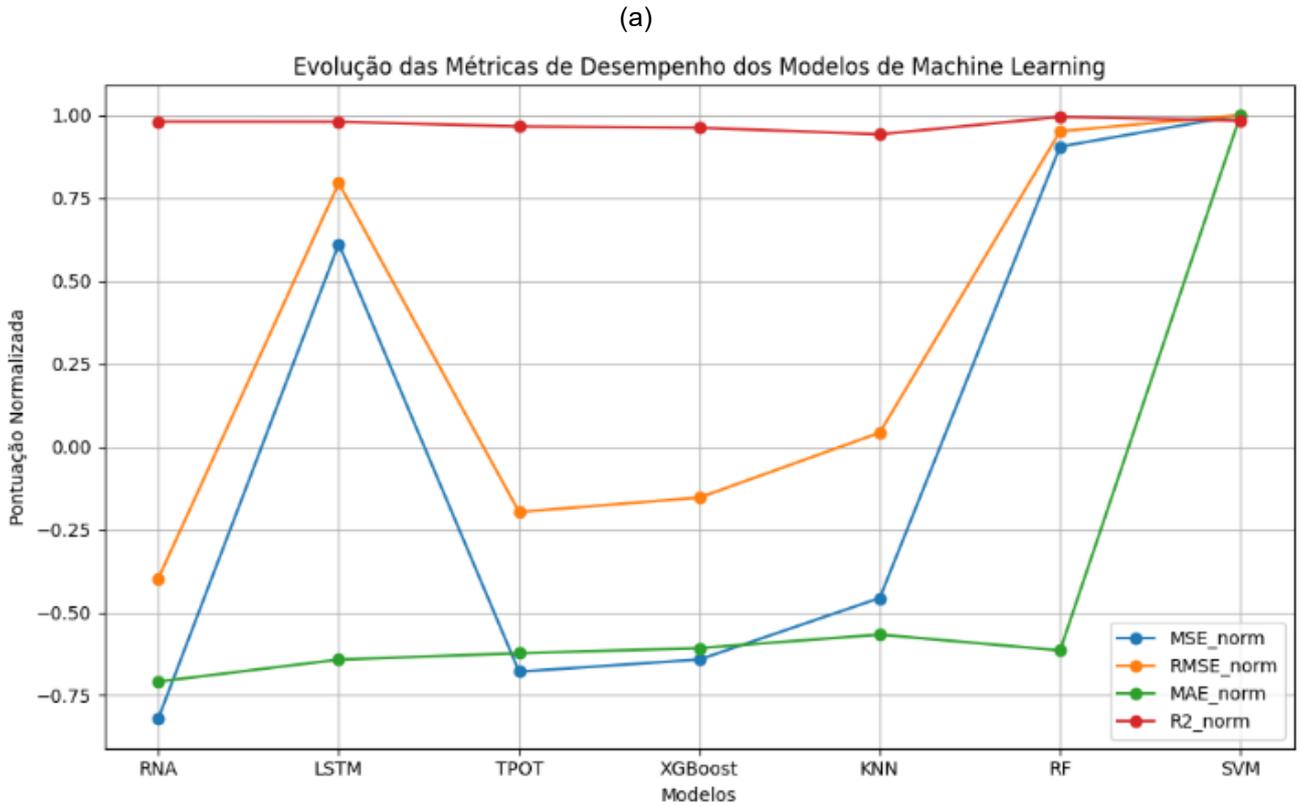
Figura 9 – Gráfico de análise de dispersão entre valores de MSE e R^2 para todos os modelos de ML avaliados.



Fonte: Autor (2024)

Um gráfico *comparativo entre as métricas normalizadas* de MSE, RMSE, MAE e R^2 (Figura 10) para os diferentes modelos também foi elaborada. RNA manteve melhor classificação em MSE e RMSE, refletindo um modelo com erros baixos e alta capacidade preditiva, consistente com o R^2 calculado (0,9903), indicando que o modelo conseguiu explicar bem a variância dos dados. O algoritmo LSTM apresentou desempenho inferior à RNA, mas com R^2 próximo (0,9901), indicando que o modelo é robusto, mas sofre maior variabilidade e/ou necessita de ajustes finos em seus parâmetros. O RF apresentou o R^2 mais alto, RMSE elevado, indicando que o modelo teve dificuldades em capturar padrões mais sutis, visto ser uma arquitetura razoavelmente de se desenhar. Já o SVM, apesar de exibir R^2 competitivo (0,9917), mas um MAE significativamente alto (0,5197), sugerindo que os erros residuais estão concentrados em casos extremos, conforme analisado anteriormente.

Figura 10 – Gráficos de (a) comparativo entre as métricas normalizadas e de (b) evolução das métricas de desempenho entre modelos para avaliação de desempenho dos modelos de ML.

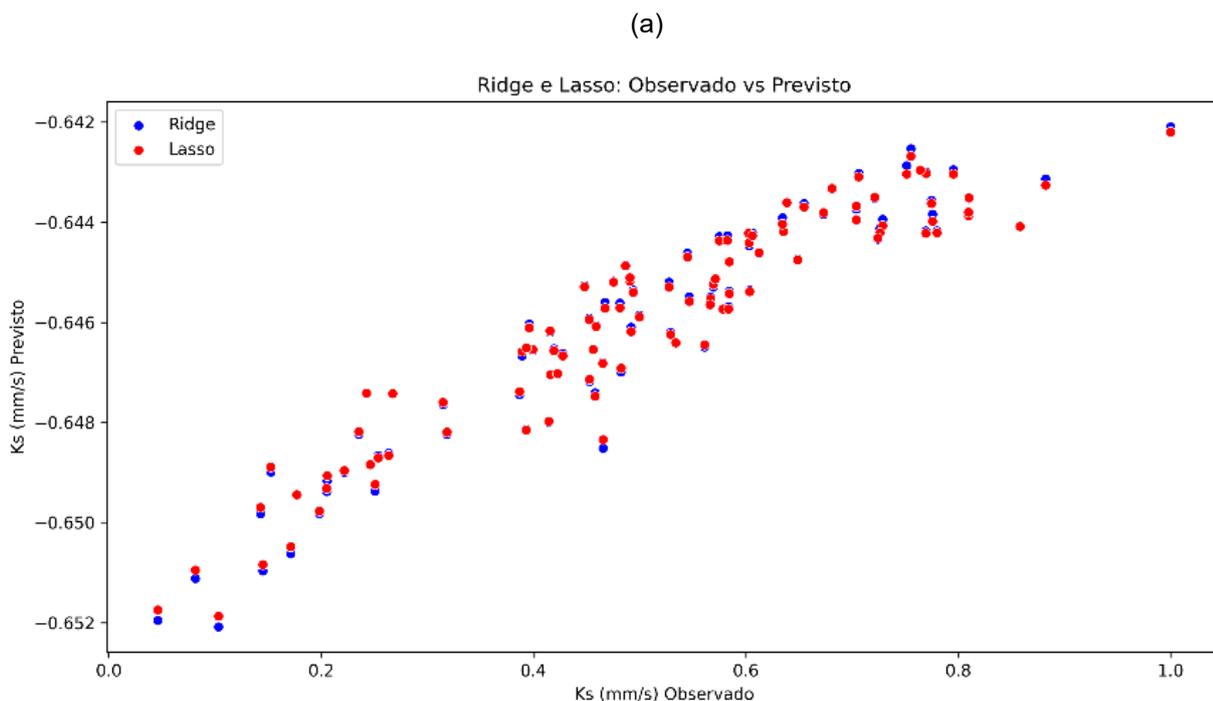


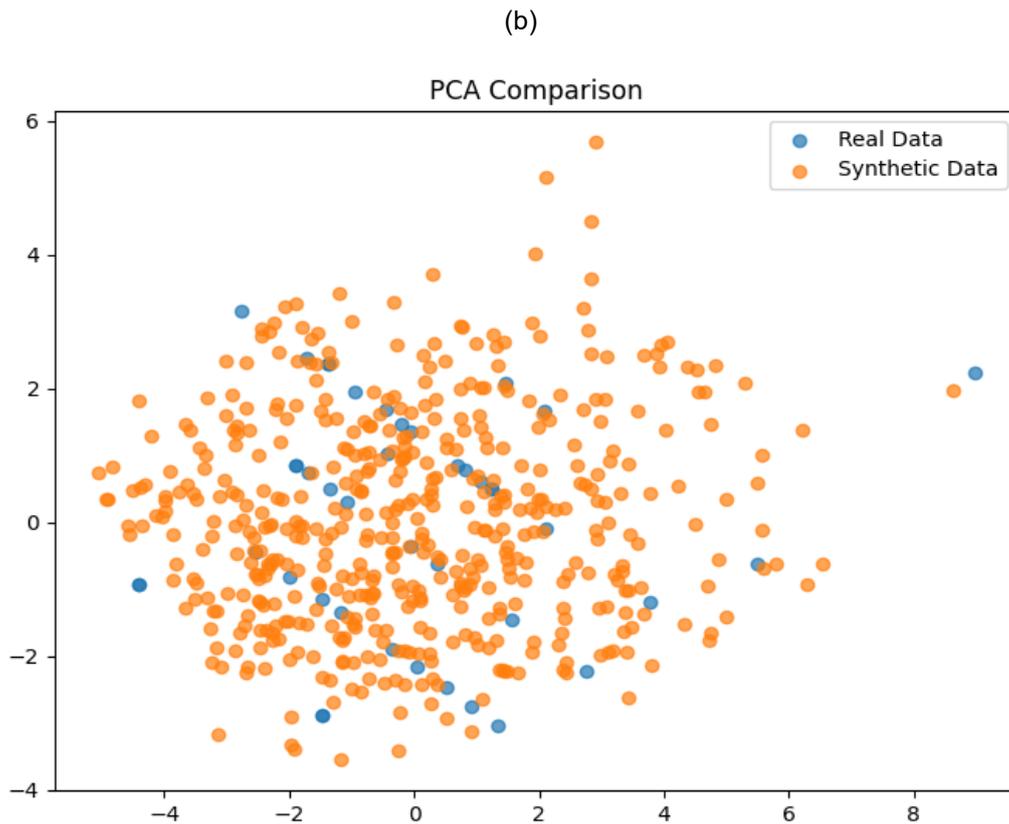
Fonte: Autor (2024)

Já o gráfico de *evolução das métricas de desempenho entre modelos* (Figura 10) detalhou a variação de métricas individuais (MSE, RMSE, MAE e R^2), confirmando tendência geral das análises anteriores, onde RNA apresentou valores consistentemente baixos de MSE, RMSE e MAE, refletindo alta eficácia.

SVM mostrou discrepância entre R^2 elevado e altos valores de MAE e RMSE, sugerindo necessidade de ajustes, baixa capacidade para interpretar os dados apresentados, e/ou overfitting em regiões específicas, embora todos os modelos tenham sido desenvolvidos implementando mecanismos de prevenção, como divisão de dados de treinamento e teste, validação cruzada (k-fold), regularização L1 Lasso, L2 Ridge (Figura 11), ajustes automáticos de hiperparâmetros (grid search, random search, bayesian optimization), *early stopping*, além da normalização dos dados antes de sua utilização. Além disso, PCA foi utilizada como mecanismo de melhoria de generalização e confirmação de correlação das variáveis. A RNA foi utilizada então para realizar demais análises da modelagem.

Figura 11 – Gráficos de regularização L1 e L2 (a) e PCA (b) de correlação da variável-alvo (K_s), durante processamento dos modelos de ML.

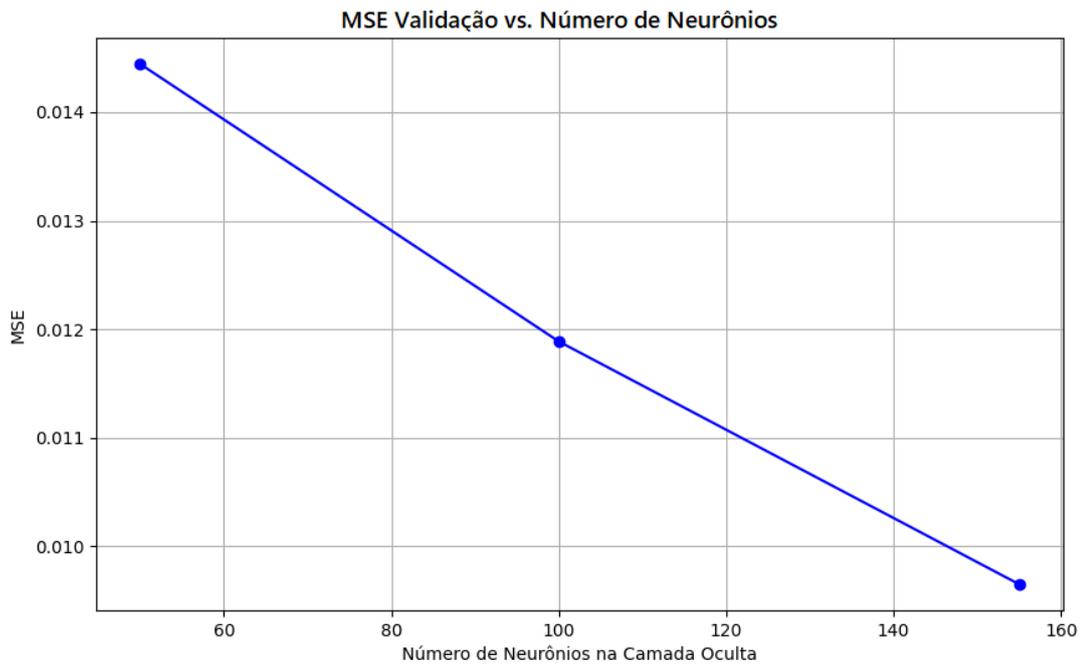
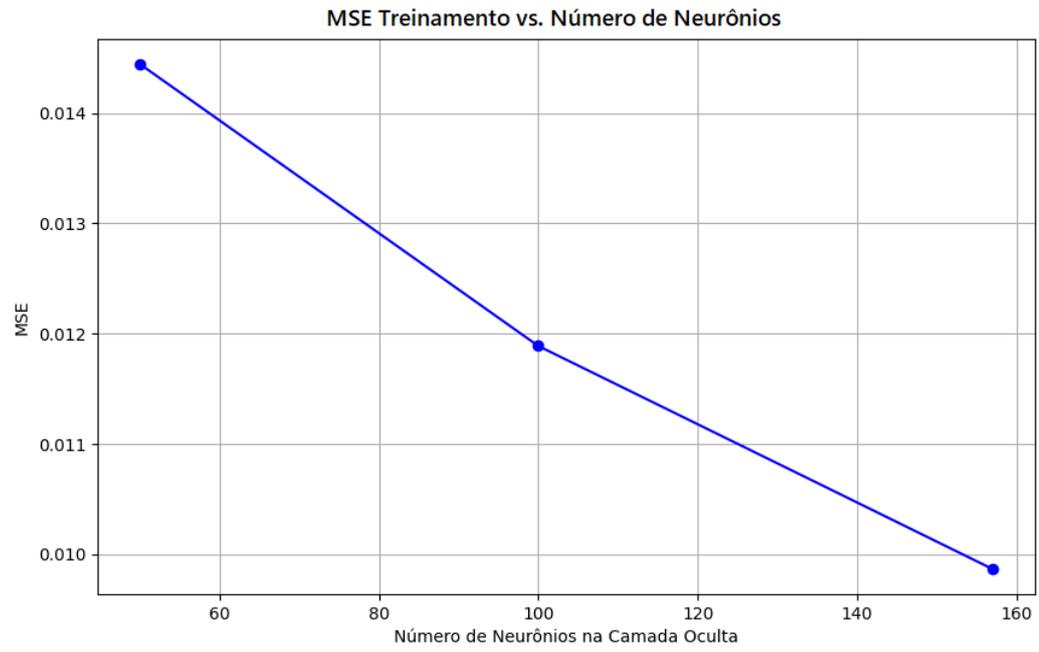


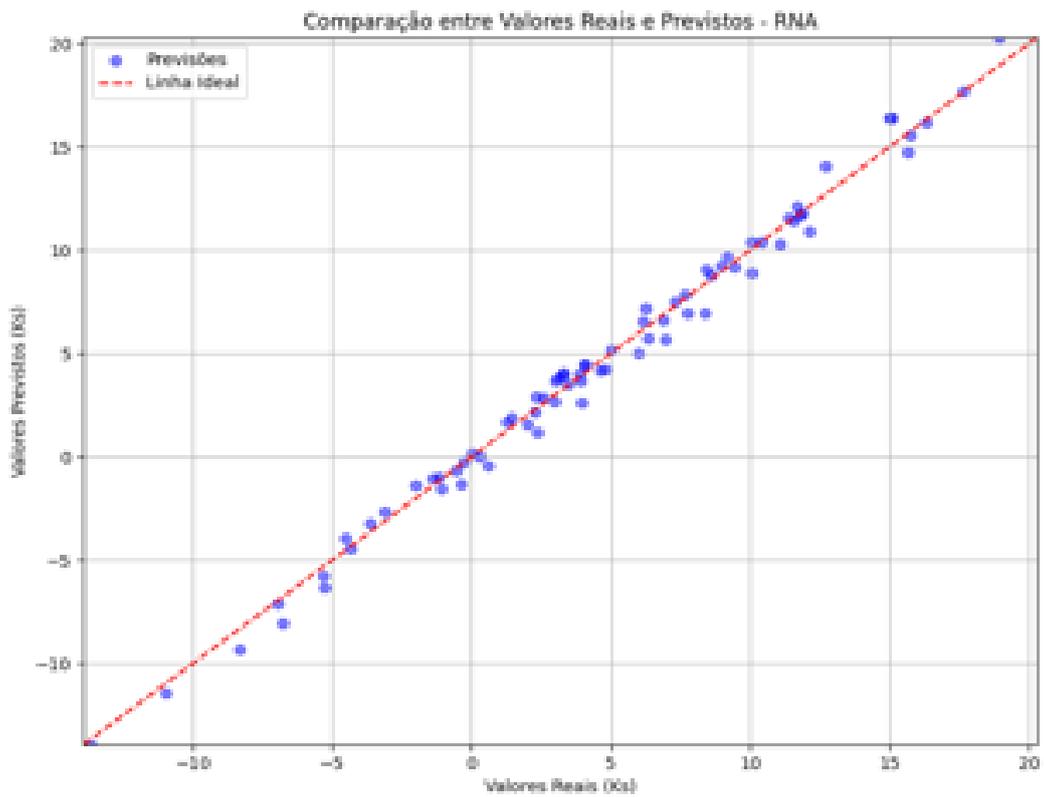
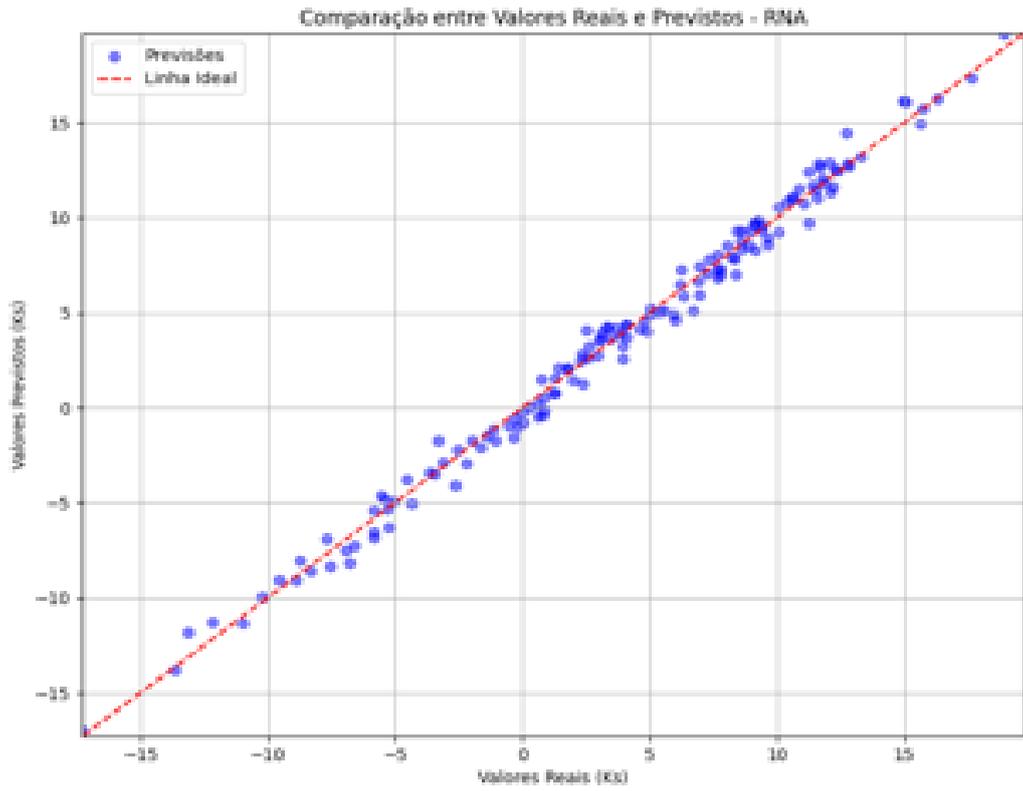


Fonte: Autor (2024)

O número de neurônios na camada escondida de uma RNA, como o MLPRegressor, da biblioteca scikit-learn Python®, utilizada neste estudo, é uma questão crítica, que pode impactar significativamente no seu desempenho, inclusive na indicação incapacidade de generalização, ou mesmo capturar padrões mais complexos nos dados, ou incorrer em overfitting. Um gráfico de *MSE versus neurônios na camada oculta* (Figura 12) avaliou o número total de neurônios necessários para se atingir o menor valor de MSE possível, atingindo um valor mínimo em torno de 157 neurônios. A validação cruzada e outros mecanismos, como já citados acima, como *early stopping*, garantindo o melhor desempenho com equilíbrio entre viés e variância, mostrando a complexidade dos dados amostrais, identificando o número ótimo de neurônios na camada oculta, resultando no melhor desempenho do modelo de RNA para estimativa de K_s .

Figura 12 – Desempenho da RNA na estimativa de K_s , avaliado pelo número de neurônios na camada escondida.





Fonte: Autor (2024)

Recentemente, modelagens com algoritmos de IA tem procurado ampliar entendimento sobre os resultados obtidos, úteis para entender interações entre variáveis preditoras e variável-alvo, validar e melhorar aplicação de modelos avaliados. Uma análise de PDP - *Partial Dependence Plots* ((Kharate et al., 2024) foi aplicada na modelagem realizada pela RNA, visando validar os resultados até então obtidos na estimação de K_s (Figura 13).

O PDP para a *areia* (Figura 13a) apresenta uma relação negativa e aproximadamente linear com a variável de saída K_s . Os valores crescentes do teor de areia influenciam em valores decrescentes de K_s , corroborando com o observado na análise SHAP, descrita a seguir, que indicou a variável areia como de impacto negativo no modelo de estimação de K_s .

O gráfico de PDP para o *silte* (Figura 13b) ressalta relação não linear, ocorrendo pico próximo a 0,15 (15% de silte), e em seguida, tendência decrescente. Isso sugere que níveis intermediários de silte podem ter um efeito positivo na condutividade hidráulica saturada, enquanto altos e baixos teores de silte tendem a diminuir o valor de K_s .

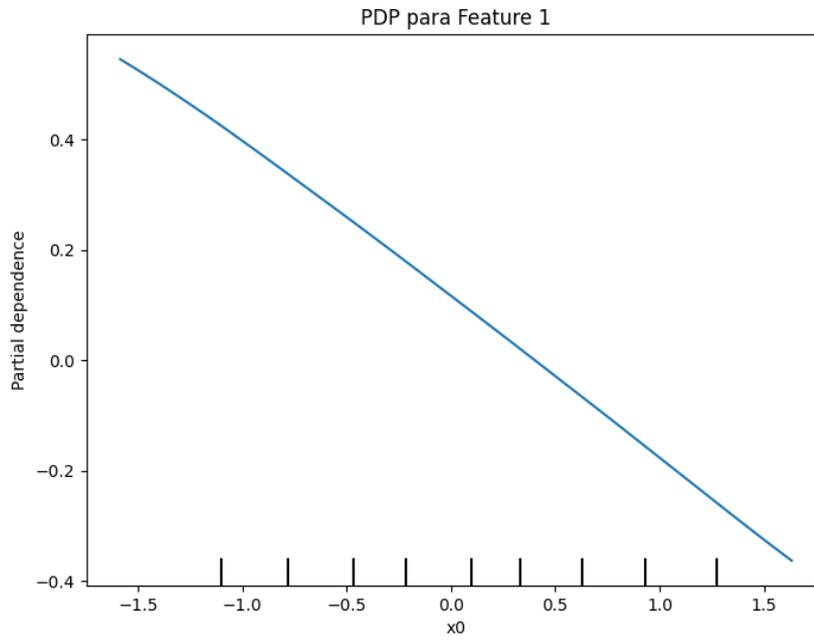
O PDP para o teor de *argila* (Figura 13c) apresenta uma relação negativa com K_s . Conforme o teor de argila aumenta, o valor da variável de saída diminui. Essa informação é consistente com a análise SHAP analisada a seguir, que apontou a argila como uma variável com impacto negativo no modelo.

Por fim, o gráfico de PDP mostra a relação entre a *densidade do solo* (Figura 13d) e a variável K_s . Foi observado uma relação positiva e aproximadamente linear entre p_b e K_s . A evolução crescente dos valores de densidade do solo é acompanhadas pelos valores de condutividade hidráulica saturada, indicando que a variável p_b mantém forte relacionamento positivo com K_s .

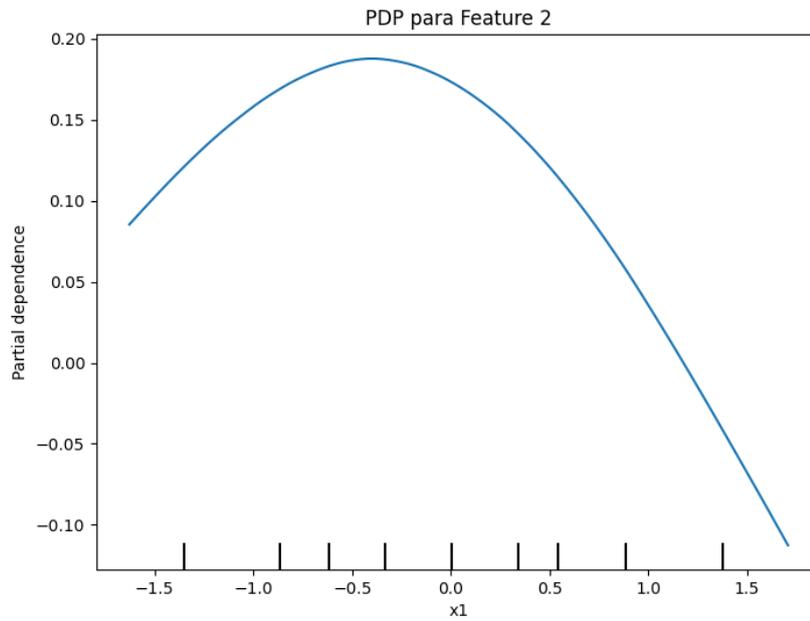
Em geral, a análise dos gráficos de PDP evidenciou informações importantes sobre o relacionamento entre as variáveis preditoras e alvo envolvidas nesta modelagem. A análise PDP complementa o entendimento da análise SHAP (que descreve a importância relativa entre as variáveis), evidenciando a natureza linear ou não linear, positiva ou negativa das relações entre as variáveis, fornecendo compreensão abrangente dos principais fatores que interferem na variável-alvo, delineando os efeitos individuais das variáveis para o desenvolvimento de modelos de pedotransferência mais robustos e interpretáveis.

Figura 13 – Análise da forma de relacionamento entre variáveis preditoras e alvo.

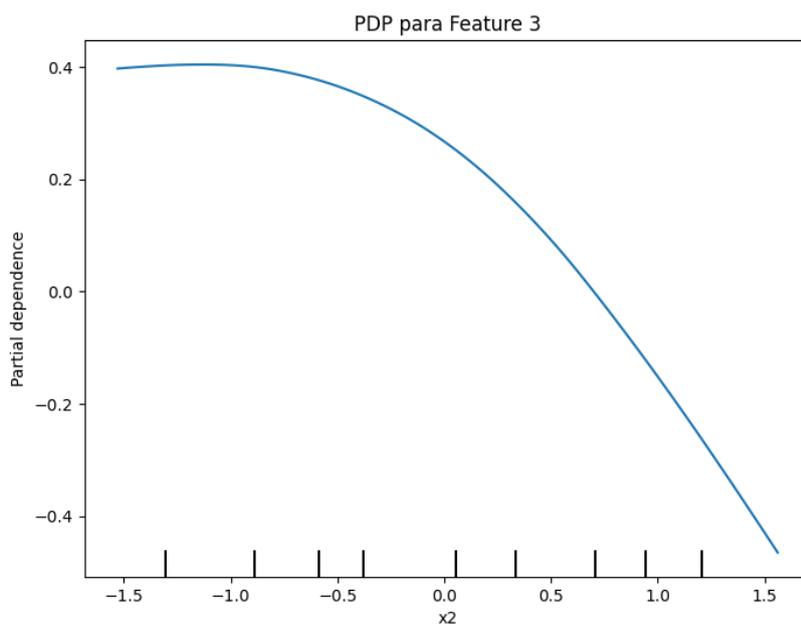
(a)



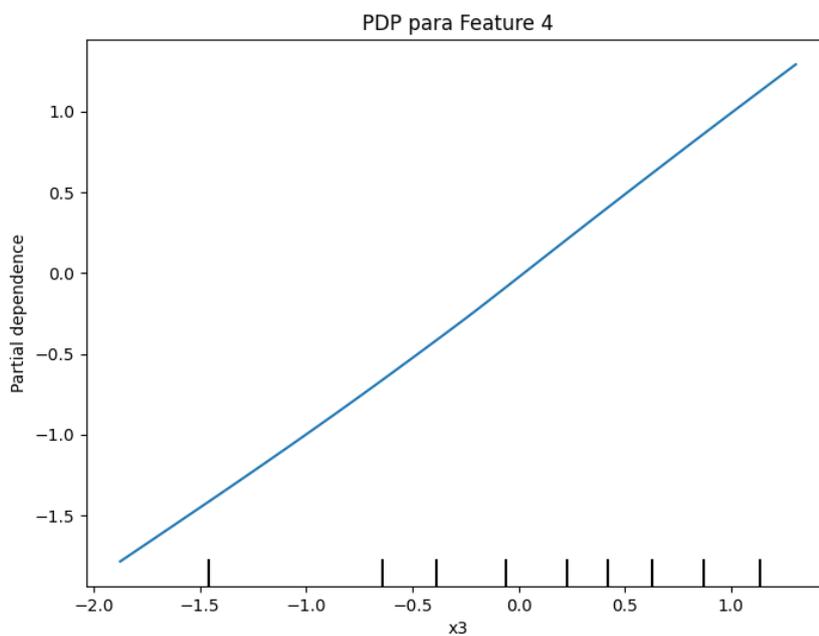
(b)



(c)



(d)



Fonte: Autor (2024)

Gráficos de importância de *features* usando SHAP - *Shapley Additive ExPlanation* (Salih et al., 2024) foram implementados, visando complementar a validação dos resultados produzidos pelo modelo neural.

O gráfico apresenta os valores SHAP (Figura 14) para diferentes variáveis de entrada do modelo, auxiliando no entendimento não apenas da direção de seus relacionamentos (positiva ou negativa), mas também da magnitude do impacto de cada característica, interpretando a contribuição de cada variável na previsão do modelo. As variáveis foram ordenadas do maior impacto médio (em valor absoluto) para o menor impacto médio no resultado do modelo predito.

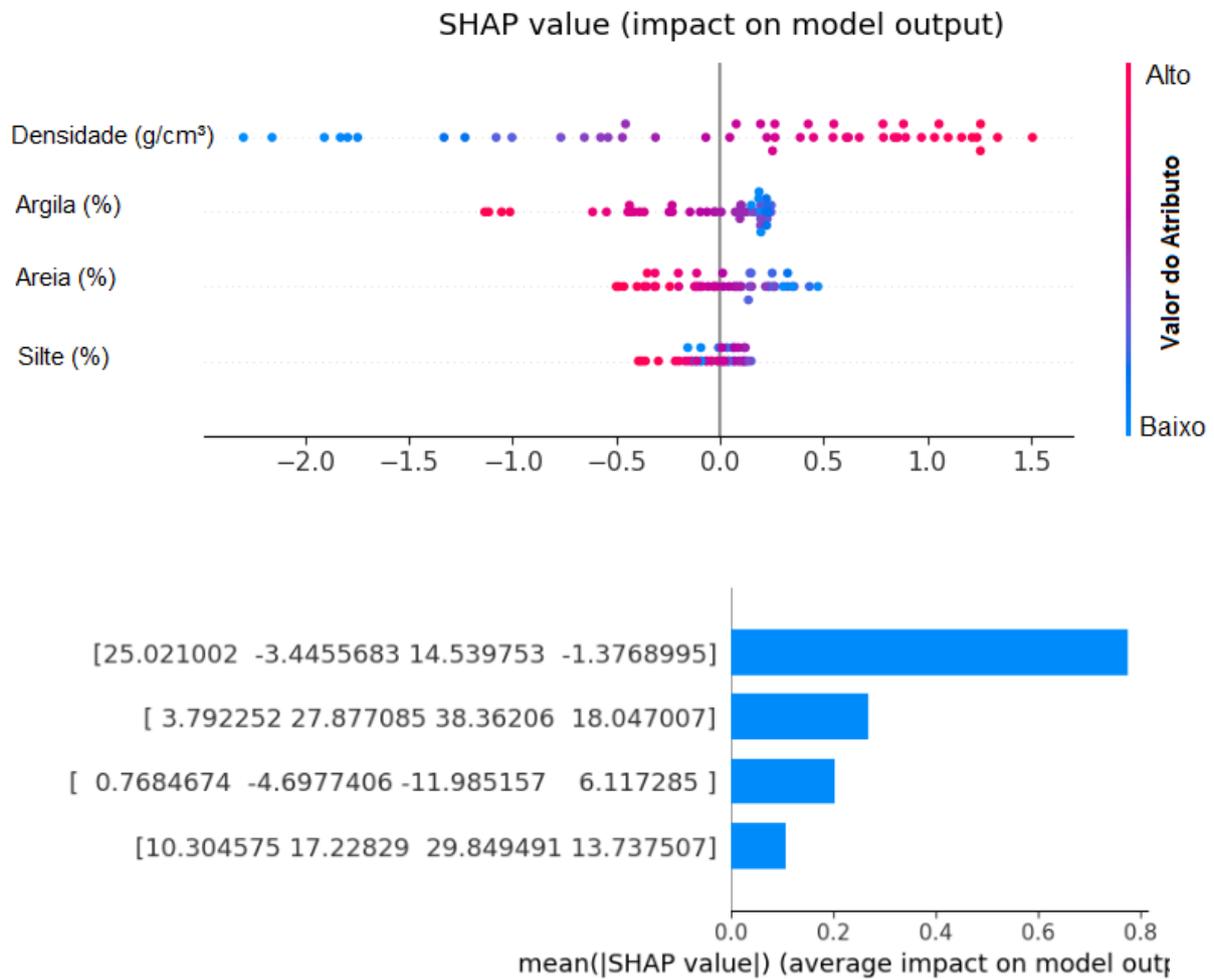
O primeiro valor ([25.021002 -3.4455683 14.539753 -1.3768995]) representa o impacto médio da variável mais influente no modelo. Esse valor é a média dos valores SHAP para essa variável, com demais valores nesta sequência, ressaltando o impacto médio das demais variáveis por sua ordem de importância. Valores positivos indicam que aumento da quantidade desta variável tende a aumentar o valor previsto pelo modelo, enquanto os valores negativos indicam que o aumento de quantidade desta variável tende a diminuir o valor previsto.

A *densidade do solo* apresentou o maior impacto absoluto nas previsões de K_s . Aumento na densidade do solo geralmente leva a intrínseco aumento no valor previsto de K_s , indicando relação direta e não linear.

A variável *argila* apresentou também importante, porém de relacionamento complexo, indicando alguns pontos de aumento em K_s , e outros de diminuição, sugerindo que a relação entre argila e K_s não seja linear, dependendo de outros fatores. Este comportamento é consistente com a natureza da textura do solo, onde solos mais argilosos geralmente apresentam menor K_s devido à compactação.

As variáveis *areia* e *silte* mostraram menor impacto em comparação com a densidade do solo e argila. A direção do impacto variou, indicando que a influência dessas variáveis pode ser mais complexa, dependendo de outras características do solo. Valores altos de areia têm impacto positivo significativo em K_s , o que é esperado, visto que solos arenosos possuem maior permeabilidade. Assim, os gráficos SHAP indicaram a densidade do solo como característica mais importante na previsão de K_s . Entretanto, a relação com as outras características e K_s é mais complexa e pode depender de interações com outras variáveis ou de condições específicas do solo.

Figura 14 – Gráficos SHAP com direção, complexidade e relacionamento entre variáveis na modelagem IA.



Fonte: Autor (2024).

Por fim, análises de matriz de correlação entre atributos principais e gráfico de índices de sensibilidade total (análise Sobol) foram avaliadas, visando investigar comportamento das variáveis preditoras em relação ao atributo-alvo, ressaltando anormalidades ou discrepâncias, ou mesmo corroborando resultados das análises anteriores.

O gráfico da matriz de correlação (Figura 15) entre atributos principais (Abba et al., 2023; Jalal et al., 2023) resalta correlação entre variáveis preditoras e alvo. Ele demonstra que argila e densidade do solo (ρ_b) mantém correlação positiva.

Areia e argila apresentam correlação negativa. Já densidade do solo (ρ_b) expressa correlação mais intensa com as demais variáveis, reforçando sua importância no modelo de predição. A correlação moderada (0.51) entre areia e silte sugere que, embora haja uma relação entre quantidade de areia e silte no solo, elas não são fortemente interdependentes.

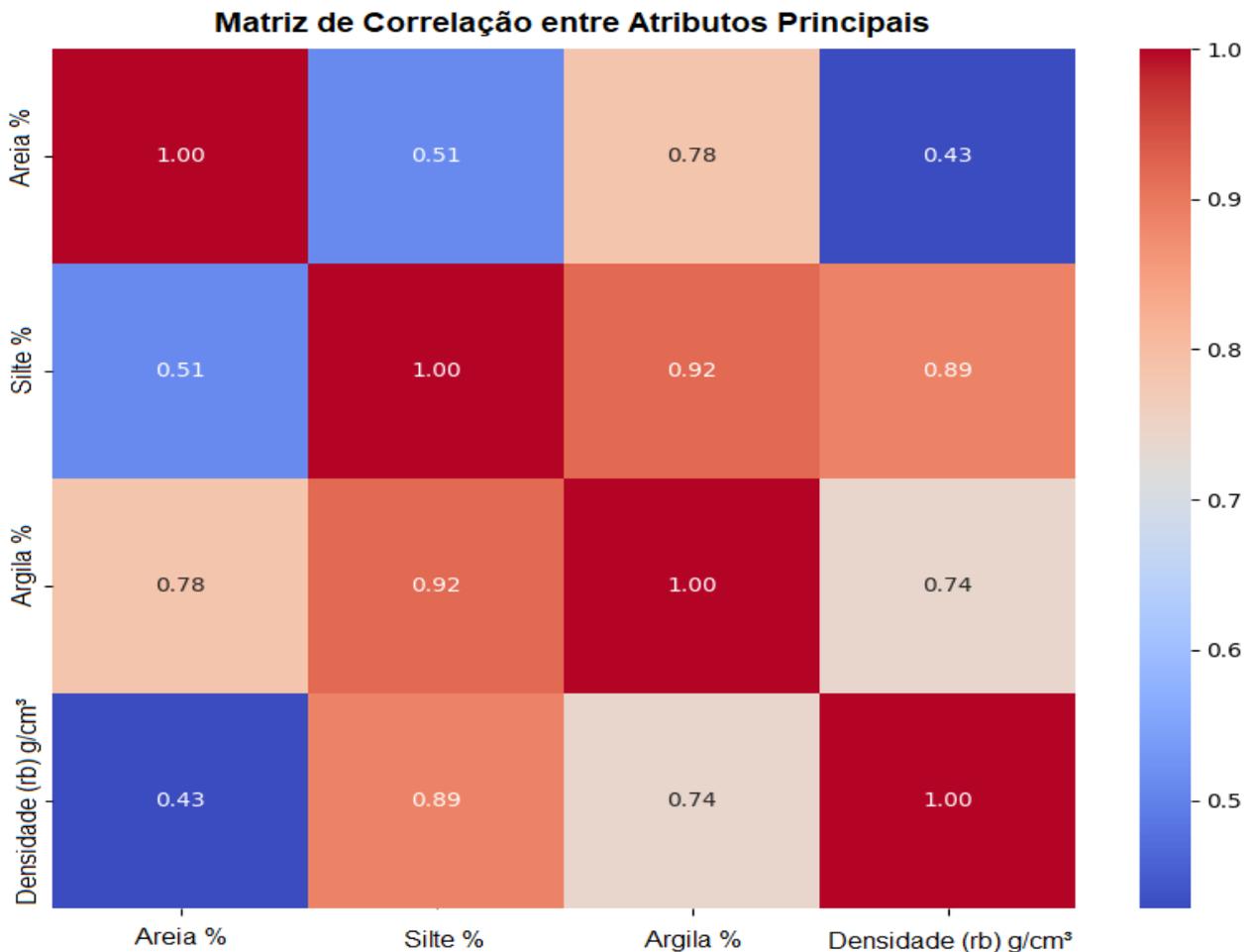
Em solos com maior teor de areia, o silte pode variar independentemente, influenciando propriedades como a permeabilidade e retenção de água. A forte correlação (0.78) entre areia e argila indica que solos com maior teor de areia tendem a apresentar também maiores teores de argila, fato relevante para a textura do solo, que pode influenciar sua capacidade de drenagem e aeração. No entanto, essa correlação pode variar, dependendo da origem do solo e de suas condições geológicas.

Já a fraca correlação (0,43) entre areia e densidade do solo indica que a quantidade de areia tem impacto moderado sobre sua densidade, possivelmente influenciado pela compactação ou pela estrutura do solo. A correlação forte (0,92) entre silte e argila estreita ligação entre elas. Isso reflete suas características semelhantes em termos de capacidade de retenção de água, influenciando diretamente a compactação e a condutividade hidráulica.

O valor de correlação entre silte e densidade do solo é forte (0,89), indicando que solos com teor de silte tendem a apresentar maior densidade, relacionado à maior compactação do solo e menor porosidade, afetando a mobilidade da água e aeração do solo. Por fim, a correlação moderada (0,74) entre argila e densidade do solo indicou que solos com maior teor de argila também tendem a ser mais compactos, podendo reduzir permeabilidade e disponibilidade de água para plantas.

Esta alta correlação entre as variáveis também ajuda a explicar os valores SHAP observados, uma vez que variáveis altamente correlacionadas tendem a ter impactos significativos no modelo, ressaltando a dificuldade de expansão por diversas técnicas no pequeno conjunto de dados originais com poda. Essas informações sobre as relações entre as variáveis de entrada são importantes para entender possíveis interações e efeitos combinados na estimativa de K_s .

Figura 15 – Gráfico de correlação entre as variáveis e seus impactos sobre o modelo de ML.



Fonte: Autor (2024)

O gráfico de índices de sensibilidade Sobol (Shreim et al., 2024) mostra a contribuição individual de cada variável de entrada para a variância total da variável de saída K_s , para variabilidade do modelo.

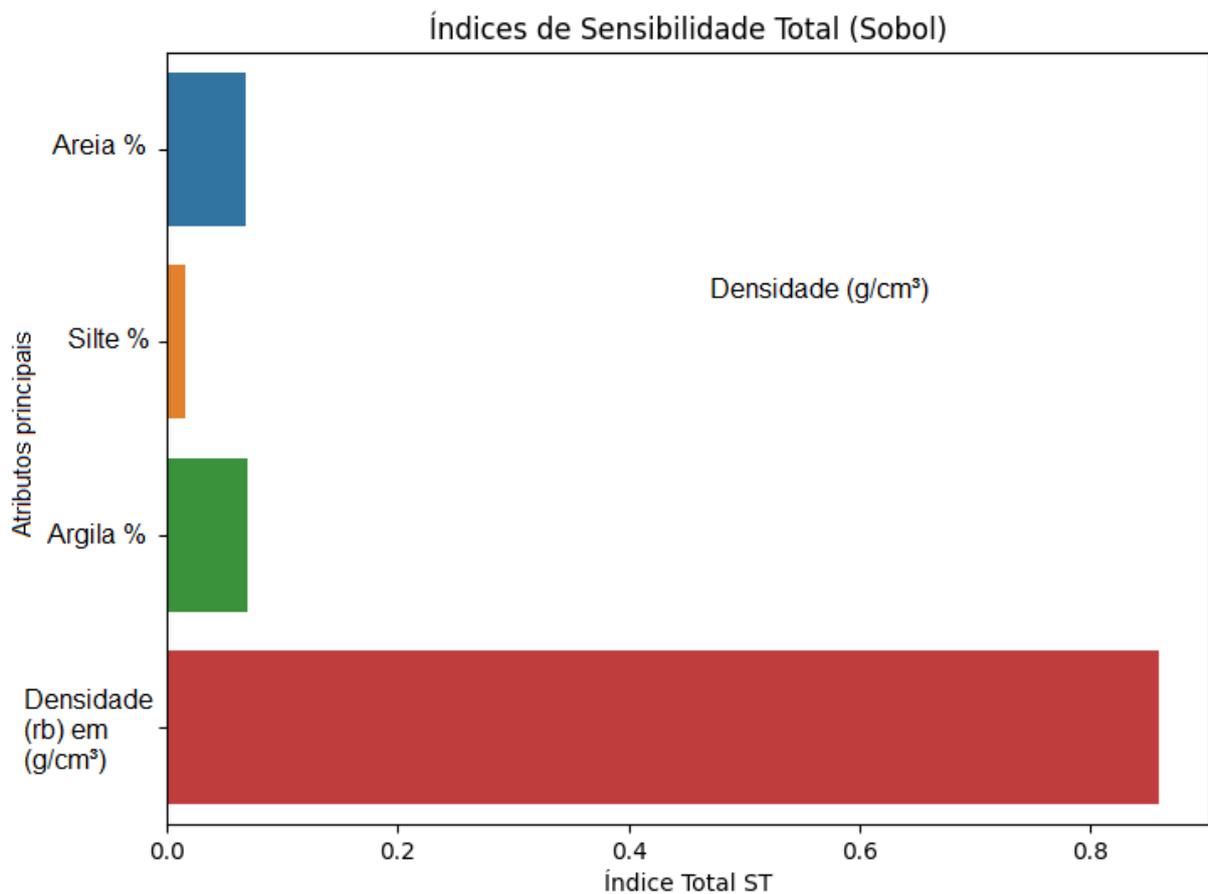
O índice de sensibilidade mais expressivo (0,8) foi obtido pela variável *densidade do solo* (ρ_b), sendo a mais influente no modelo de predição. Este cenário corrobora estudos na ciência do solo sobre sua relevância, pois mantém relação direta com propriedades físicas do solo, como compactação, porosidade, retenção de água e capacidade de drenagem.

Estes fatores são determinantes para a condutividade hidráulica saturada (K_s). A *argila* exibe o segundo maior índice de sensibilidade, sugerindo que a proporção de argila no solo

é importante, mas secundária, para a predição. A argila afeta K_s , influenciando a textura do solo, aumentando a retenção de água e reduzindo a permeabilidade. A variável *areia* apresenta menor impacto, indicando que sua influência na variabilidade do modelo é moderada, refletindo seu papel em aumentar a permeabilidade do solo.

Por fim, a *silte* apresentou o menor índice de sensibilidade total, sendo a variável menos influente no modelo, com pouca contribuição para a predição de K_s .

Figura 16 – Contribuição das variáveis predictoras para a variabilidade da modelagem IA.



Fonte: Autor (2024)

Como proposta de produto resultante da modelagem da condutividade hidráulica saturada por modelos de inteligência artificial, avaliados neste estudo, uma equação de função de pedotransferência (FPT) para estimação genérica do melhor modelo selecionado foi desenvolvido, com alvo no conjunto de dados experimentais originais, ou no expandido pela GAN.

A FPT foi definida como um modelo simples, com duas camadas lineares e ajustes não lineares, com regularização (validação cruzada), equilibrando capacidade preditiva e interpretabilidade. Na simulação da predição de K_s com RNA (Figura 17), a equação de FPT de ótimo desempenho, com R^2 explicando mais de 95% (0,9539) da variância dos dados, e erro médio absoluto (MAE) baixo (0,0100), confirmando alta precisão, com MSE médio de 0,0001.

O modelo proposto é uma abordagem inicial para funções de pedotransferência, aliando alta precisão com baixo custo computacional.

A equação do modelo de FPT foi definida como

$$K_s = |1.0 * X + 0.1| * 0.8 + 0.11$$

Onde:

- K_s é a variável de saída, representando a condutividade hidráulica saturada.
- X é a variável de entrada normalizada, no intervalo [0, 1].

A equação contém duas camadas ocultas, sendo:

Camada 1 (Ativação ReLU): $\text{layer1} = \text{np.maximum}(0, X * 1.0 + 0.1)$;

Camada 2 (Função não linear): $\text{layer2} = \text{np.abs}(\text{layer1} * 0.8 + 0.11)$.

Observações:

A **camada 1** aplica uma ativação ReLU (*Rectified Linear Unit*) à combinação linear da entrada X e do intercepto 0.1. A saída dessa camada será 0, quando a combinação linear for negativa, e igual à combinação linear, quando essa for positiva;

A **camada 2** aplica uma função não linear, usando o valor absoluto da combinação linear da saída da primeira camada (layer1) com o coeficiente 0.8 e o intercepto 0.11.

A representação geral da FPT, considerando as variáveis físicas estimadas pelo modelo neural, considerando as saídas de suas camadas, seria:

$$K_s = 1.0 \cdot (w_1 \cdot \text{Areia} + w_2 \cdot \text{Argila} + w_3 \cdot \text{Silte} + w_4 \cdot \rho_b) + 0.1 \cdot 0.8 + 0.11$$

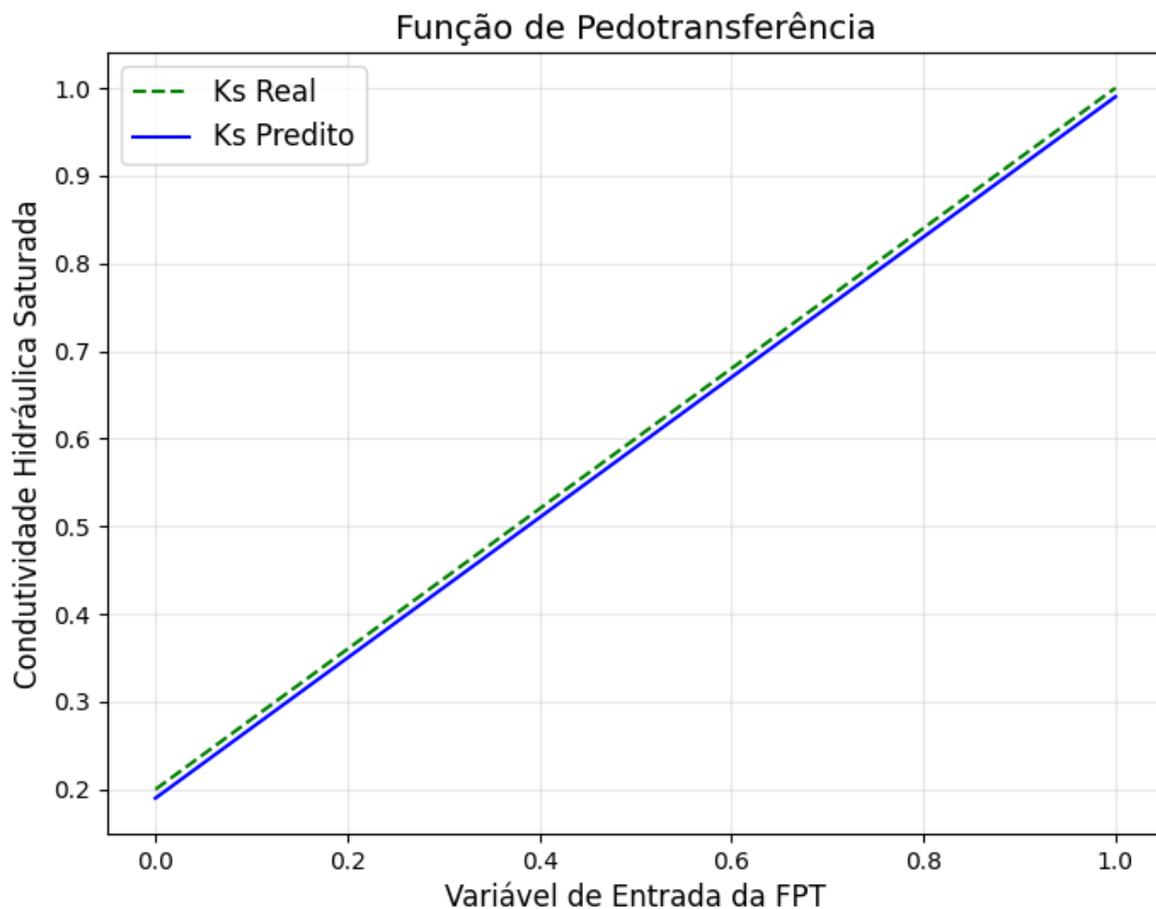
Onde:

w_i como fator de calibração do modelo para diferentes tipos de solo.

A representação específica para os valores estimados pela RNA, considerando o contexto acima:

$$K_s = 0.45 \cdot \text{Areia} - 0.30 \cdot \text{Argila} + 0.20 \cdot \text{Silte} - 0.40 \cdot \text{Densidade do solo} + 0.1 \cdot 0.8 + 0.11$$

Figura 17 – Simulação da equação de FPT com dados de K_s preditos pelo modelo da RNA.



Fonte: Autor (2024)

5.4 CONCLUSÕES

O pequeno conjunto de dados experimentais originais representou um desafio considerável para modelagem da condutividade hidráulica saturada por algoritmos de IA limitando a capacidade de generalização e replicação de padrões espaciais.

A análise por quatro técnicas de expansão de dados demonstrou que nenhuma conseguiu dados sintéticos com distribuições uniformes e adequadas. A análise dos dados expandidos por GAN, mesmo com suavização de *outliers*, mostrou-se eficaz em promover distribuições mais uniformes e adequadas para modelagem de aprendizado de máquina, sem, entretanto, resolver completamente a heterogeneidade intrínseca dos dados experimentais, ressaltando que pontos extremos, apesar de suavizados, mantêm relevância para fenômenos significativos, fundamentais para a compreensão dos processos físicos do solo, devendo ser previamente estudados, antes de sua exclusão.

Embora outros fatores possam ter influenciado na performance dos modelos de IA, nem todos se mostraram adequados para predição de variáveis complexas, como as ambientais, necessitando de premissas adequadas para escolha do modelo a ser aplicado, visando garantir maximização da estimação pela modelagem.

As análises dos resultados alcançados neste estudo indicaram que um número mínimo de variáveis, além de sua importância e complexidade afetam o sucesso da modelagem, influenciando no custo computacional, assim como na acurácia obtida. Em vários estudos pesquisados na literatura para predição de condutividade hidráulica saturada, uma média de 6 variáveis preditoras têm sido utilizadas.

A implementação de análises SHAP e Sobol complementaram e validaram a compreensão sobre contribuição de cada variável preditora, destacando densidade do solo como a mais influente neste trabalho (impacto global de 0,8 no índice Sobol e valores SHAP elevados), apresentando relação direta, linear positiva com K_s . Areia e argila mostraram impactos significativos, secundários, com comportamentos complexos e não lineares. A inclusão de mais variáveis mais impactantes poderiam ser investigadas em futuros estudos, permitindo pesquisadores potencializarem seus efeitos na modelagem com algoritmos de ML.

As análises PDP corroboraram os insights obtidos por SHAP, permitindo explorar a natureza das relações entre variáveis preditoras e K_s (lineares ou complexas). A matriz de correlação adicionou informações importantes sobre relações entre variáveis argila, silte e

densidade do solo, indicando interdependências que podem afetar as propriedades hidráulicas do solo.

A pesquisa corroborou ainda que algoritmos neurais funcionam como modelos caixa-preta, devendo ser auxiliados de técnicas de análises de dados e de resultados preditos mais robustos, como forma de garantir confiabilidade e comprovar precisão da modelagem.

O modelo de RNA mostrou robustez e desempenho para atuar com dados de relações complexas e heterogêneas. Os demais modelos foram afetados por fatores diversos, onde sensibilidade, a distribuição não normalizada afetou a LSTM. RF demonstrou robustez a outliers, mas não conseguiu capturar padrões sutis dos dados, refletindo seu comportamento como modelo de aprendizado ensemble. SVM e KNN sofreram desafios agressivos, com altos valores de erro (MAE e RMSE), mostrando-se inadequados para estes tipos de dados. Ajustes adicionais de hiperparâmetros poderiam ser investigados para melhorar desempenho de modelos como XGBoost e RF.

Embora métodos baseados em RNA e a aprendizagem profunda (LSTM) apresentem resultados promissores, exigem frequentemente recursos computacionais substanciais. Outro fator importante é a definição de arquitetura e calibração dos modelos. A ocorrência de sobre-ajustamento dos resultados também é frequente, especialmente quando se trata de conjuntos de dados pequenos ou desequilibrados, exigindo implementação de mecanismos de prevenção e tratamento de inconsistências. A comprovação de resultados é essencial para validação da pesquisa, satisfeita com integração de técnicas explicativas, como SHAP e Sobol, proporcionando análise aprofundada dos fatores determinantes na modelagem de dados.

Apesar de processos ambientais serem representados por variáveis que possuem alta dimensionalidade, complexidade, além de dependência temporal e espacial, especialmente quando há pouca disponibilidade de conjunto amostral, apresentando qualidade moderada, com ocorrência de dados esparsos, *outliers*, e presença de ruídos, tornando-se desafios e limitações para sua modelagem, este estudo procurou demonstrar que ferramentas de inteligência artificial, especificamente algoritmos de aprendizado de máquina, se adequadamente aplicados, podem ser poderosas ferramentas de auxílio para a pesquisa na área ambiental.

O trabalho também procurou demonstrar que procedimentos computacionais e estatísticos, corretamente aplicados, sequenciando etapas da modelagem, podem servir de diretriz para selecionar e tratar arquivos de entrada, tratamento e expansão de base de

dados, indicação e avaliação de modelos de IA apropriados, assim como análise robusta dos resultados obtidos. Ele apresenta um conjunto de premissas e procedimentos ideais a serem adotados em investigações semelhantes, servindo de guia para modelagem de dados ambientais com modelos de IA inclusive inovando, como a utilização da modelagem para determinação de equações experimentais do solo, contribuindo para o desenvolvimento de modelos de pedotransferência interpretáveis e robustos.

Os resultados obtidos neste trabalho disponibilizam valioso acervo de informações para complementar pesquisas experimentais e investigações futuras, com potencial para melhorar o manejo e a sustentabilidade do uso da água no solo.

REFERÊNCIAS

- ABBA, S. I.; BENAIFI, M.; ALJUNDI, I. H. Intelligent process optimisation based on cutting-edge emotional learning for performance evaluation of NF/RO of seawater desalination plant. **Desalination**, v. 550, p. 116376, mar. 2023.
- AHMADISHARAF, A. et al. Representative Sample Size for Estimating Saturated Hydraulic Conductivity via Machine Learning: A Proof-Of-Concept Study. **Water Resources Research**, v. 60, n. 8, p. e2023WR036783, ago. 2024.
- ALBALASMEH, A. et al. Artificial neural network optimization to predict saturated hydraulic conductivity in arid and semi-arid regions. **CATENA**, v. 217, p. 106459, out. 2022.
- CHANG, F., YANG, W., WANG, S., YIN, L., & DENG, X. Long-term optimization of agronomic practices increases water storage capacity and available water in soil. **Agronomy**, 14(10), 2286, 2024.
- CHEN, B.; LI, J.; JIN, Y. Deep Learning for Feature-Level Data Fusion: Higher Resolution Reconstruction of Historical Landsat Archive. **Remote Sensing**, v. 13, n. 2, p. 167, 6 jan. 2021.
- CHOUDHARY, C. S. et al. Mechanisms of allelopathic interactions for sustainable weed management. **Rhizosphere**, v. 25, p. 100667, mar. 2023.
- DE RIJK, V. et al. **Predicting Saturated Hydraulic Conductivity from Particle Size Distributions Using Machine Learning**. SSRN, 2024. Disponível em: <<https://www.ssrn.com/abstract=4862913>>. Acesso em: 16 nov. 2024
- FAE, G. S., KEMANIAN, A. R., ROTH, G. W., WHITE, C., & WATSON, J. E. Soybean yield in relation to environmental and soil properties. **European Journal of Agronomy**, 118, 126070, 2020.
- FAOUZI, E. et al. Spatial mapping of hydrologic soil groups using machine learning in the Mediterranean region. **CATENA**, v. 232, p. 107364, nov. 2023.
- GONZALES-INCA, C. et al. Geospatial Artificial Intelligence (GeoAI) in the Integrated Hydrological and Fluvial Systems Modeling: Review of Current Applications and Trends. **Water**, v. 14, n. 14, p. 2211, 13 jul. 2022.

- GUO, H. et al. Robust Attentive Deep Neural Network for Detecting GAN-Generated Faces. **IEEE Access**, v. 10, p. 32574–32583, 2022.
- GUPTA, D.; DAS, A.; MITRA, S. Role of modeling and artificial intelligence in process parameter optimization of biochar: A review. **Bioresource Technology**, v. 390, p. 129792, dez. 2023.
- HE, L., DUAN, X., DING, S., CAI, C., & DENG, Y. Construction and evaluation of pedotransfer functions for saturated hydraulic conductivity in the granite red soil regions of southern China. **Journal of Hydrology: Regional Studies**, v. 50, p. 101539, dez. 2023.
- JALAL, F. E., IQBAL, M., ALI KHAN, M., SALAMI, B. A., ULLAH, S., KHAN, H., & NABIL, M. Indirect Estimation of Swelling Pressure of Expansive Soil: GEP versus MEP Modelling. **Advances in Materials Science and Engineering**, v. 2023, p. 1–25, 23 jan. 2023.
- JEBRAEELI, V., JIANG, B., KRIM, H., & CANSEVER, D. **Generative Expansion of Small Datasets: An Expansive Graph Approach**. arXiv, 1 out. 2024. Disponível em: <<http://arxiv.org/abs/2406.17238>>. Acesso em: 10 dez. 2024.
- JOZDANI, S., CHEN, D., POULIOT, D., & ALAN JOHNSON, B. A review and meta-analysis of Generative Adversarial Networks and their applications in remote sensing. **International Journal of Applied Earth Observation and Geoinformation**, v. 108, p. 102734, abr. 2022.
- KANG, W., TIAN, J., REEMT BOGENA, H., LAI, Y., XUE, D., & HE, C. Soil moisture observations and machine learning reveal preferential flow mechanisms in the Qilian Mountains. **Geoderma**, v. 438, p. 116626, out. 2023.
- KARUNATHILAKE, E. M. B. M., LE, A. T., HEO, S., CHUNG, Y. S., & MANSOOR, S. The Path to Smart Farming: Innovations and Opportunities in Precision Agriculture. **Agriculture**, v. 13, n. 8, p. 1593, 11 ago. 2023.
- KEÏTA, A., ZOROM, M., FAYE, M. D., DAMBA, D. D., KONATE, Y., HAYDE, L. G., & LIDON, B. Achieving Real-World Saturated Hydraulic Conductivity: Practical and Theoretical Findings from Using an Exponential One-Phase Decay Model. **Hydrology**, v. 10, n. 12, p. 235, 9 dez. 2023.
- KHARATE, N., ANERAO, P., KULKARNI, A., & ABDULLAH, M. Explainable AI Techniques for Comprehensive Analysis of the Relationship between Process Parameters and Material Properties in FDM-Based 3D-Printed Biocomposites. **Journal of Manufacturing and Materials Processing**, v. 8, n. 4, p. 171, 6 ago. 2024.
- KIRAN, A.; KUMAR, S. S. A Methodology and an Empirical Analysis to Determine the Most Suitable Synthetic Data Generator. **IEEE Access**, v. 12, p. 12209–12228, 2024.
- KORKMAZ, M. SoC estimation of lithium-ion batteries based on machine learning techniques: A filtered approach. **Journal of Energy Storage**, v. 72, p. 108268, nov. 2023.
- KUMAR, V., KEDAM, N., SHARMA, K. V., MEHTA, D. J., & CALOIERO, T. Advanced Machine Learning Techniques to Improve Hydrological Prediction: A Comparative Analysis of Streamflow Prediction Models. **Water**, v. 15, n. 14, p. 2572, 13 jul. 2023.
- LI, D., JIANG, Z., TIAN, K., JI, R. Prediction of hydraulic conductivity of sodium bentonite gels by machine learning approaches. **Environmental Geotechnics**, e-issn 2051-803x, (2023), <https://doi.org/10.1680/jenge.22.00181>.
- LIU, H., LI, B., WU, H., LIANG, H., HUANG, Y., LI, Y., GHANEM, B., & ZHENG, Y. Combating Mode Collapse via Offline Manifold Entropy Estimation. **Proceedings of the AAAI Conference on Artificial Intelligence**, v. 37, n. 7, p. 8834–8842, 26 jun. 2023.

LIU, X. "Deep learning in marketing: a review and research agenda", SUDHIR, K. AND TOUBIA, O. (ed.) artificial intelligence in marketing (review of marketing research, vol. 20), **Emerald Publishing Limited**, leeds, pp. 239-271, (2023).

LÜ, X., NURMEMET I., XIAO, S., ZHAO, J., YU, X., AILI, Y., LI, S. Spatial-temporal simulation and prediction of root zone soil moisture based on hydrus-1d and cnn-lstm-attention in the yutian oasis, southern xinjiang, china, **Pedosphere**, 2024, issn 1002-0160, (2024). <https://doi.org/10.1016/j.pedsph.2024.10.008>.

MAHSA H. KASHANI, MOHAMMAD ALI GHORBANI, MAHMOOD SHAHABI, SUJAY RAGHAVENDRA NAGANNA, LAMINE DIOP. Multiple ai model integration strategy—application to saturated hydraulic conductivity prediction from easily available soil properties. **Soil and Tillage Research**, volume 196, 104449, issn 0167-1987, 2020.

MEHMOOD R, BASHIR R, & GIRI KJ Deep generative models: a review. department of computer science, islamic university of science & technology, kashmir, jammu and kashmir, india. **Indian Journal of Science and Technology**, 16(7), 460–467, 2023.

MOZAFFARI, H.; MOOSAVI, A. A.; NEMATOLLAHI, M. A. Predicting saturated and near-saturated hydraulic conductivity using artificial neural network and multiple linear regression in calcareous soils. **PLOS ONE**, v. 19, n. 1, p. e0296933, 10 jan. 2024.

MUÑOZ-CARPENA, R., CARMONA-CABRERO, A., YU, Z., FOX, G., & BATELAAN, O. Convergence of mechanistic modeling and artificial intelligence in hydrologic science and engineering. **PLOS Water**, v. 2, n. 8, p. e0000059, 7 ago. 2023.

PULLA, S. T.; YASARER, H.; YARBROUGH, L. D. Synthetic Time Series Data in Groundwater Analytics: Challenges, Insights, and Applications. **Water**, v. 16, n. 7, p. 949, 25 mar. 2024.

REZAEI, M., MOUSAVI, S. R., RAHMANI, A., ZERAATPISHEH, M., RAHMATI, M., PAKPARVAR, M., JAHANDIDEH MAHJENABADI, V. A., SEUNTJENS, P., & CORNELIS, W. Incorporating machine learning models and remote sensing to assess the spatial distribution of saturated hydraulic conductivity in a light-textured soil. **Computers and Electronics in Agriculture**, v. 209, p. 107821, jun. 2023.

ROBERT, C. P.; ROUSSEAU, J. A special issue on Bayesian inference: challenges, perspectives and prospects. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 381, n. 2247, p. 20220155, 15 de maio 2023.

SALIH, A., RAISI-ESTABRAGH, Z., GALAZZO, I. B., RADEVA, P., PETERSEN, S. E., MENEGAZ, G., & LEKADIR, K. Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME. **Advanced Intelligent Systems**, p. 2400304, 27 jun. 2024.

SHAN, W.; LIU, J.; ZHOU, J. Integrated method for intelligent structural design of steel frames based on optimization and machine learning algorithm. **Engineering Structures**, v. 284, p. 115980, jun. 2023.

SHI, S., TSE, R., LUO, W., D'ADDONA, S., & PAU, G. Machine learning-driven credit risk: a systemic review. **Neural Computing and Applications**, v. 34, n. 17, p. 14327–14339, set. 2022.

SHREIM, H.; GIZZINI, A. K.; GHANDOUR, A. J. **Trainable Noise Model as an Explainable Artificial Intelligence Evaluation Method: Application on Sobol for Remote Sensing Image Segmentation**. ECRS 2023. **Anais**. Em: ECRS 2023. MDPI, 6 nov. 2023. Disponível em: <<https://www.mdpi.com/2673-4931/29/1/49>>. Acesso em: 12 dez. 2024.

- SINGH, V. K. et al. Novel Genetic Algorithm (GA) based hybrid machine learning-pedotransfer Function (ML-PTF) for prediction of spatial pattern of saturated hydraulic conductivity. **Engineering Applications of Computational Fluid Mechanics**, v. 16, n. 1, p. 1082–1099, 31 dez. 2022.
- SLATER, L. J., ARNAL, L., BOUCHER, M.-A., CHANG, A. Y.-Y., MOULDS, S., MURPHY, C., NEARING, G., SHALEV, G., SHEN, C., SPEIGHT, L., VILLARINI, G., WILBY, R. L., WOOD, A., AND ZAPPA, M.: Hybrid forecasting: blending climate predictions with ai models, hydro. **Earth Syst. Sci.**, 27, 1865–1889, 2023.
- TAN, W., LIU, P., LI, X., XU, S., CHEN, Y., & YANG, J. Segmentation of lung airways based on deep learning methods. **IET Image Processing**, v. 16, n. 5, p. 1444–1456, abr. 2022.
- TARWIDI, D., PUDJAPRASETYA, S. R., ADYTIA, D., & APRI, M. TARWIDI, D. et al. An optimized XGBoost-based machine learning method for predicting wave run-up on a sloping beach. **MethodsX**, v. 10, p. 102119, 2023.
- TREJO-ALONSO, J., FUENTES, S., MORALES-DURAN, N., & CHAVEZ, C. Saturated Hydraulic Conductivity Estimation Using Artificial Neural Networks. **Water**, v. 13, n. 5, p. 705, 5 mar. 2021.
- VELOSO, M. F.; RODRIGUES, L. N.; FILHO, E. I. F. Evaluation of machine learning algorithms in the prediction of hydraulic conductivity and soil moisture at the Brazilian Savannah. **Geoderma Regional**, v. 30, p. e00569, set. 2022.
- VON KROGH, G.; ROBERSON, Q.; GRUBER, M. Recognizing and Utilizing Novel Research Opportunities with Artificial Intelligence. **Academy of Management Journal**, v. 66, n. 2, p. 367–373, abr. 2023.
- WANG, J., ZHEN, J., HU, W., CHEN, S., LIZAGA, I., ZERAATPISHEH, M., & YANG, X. Remote sensing of soil degradation: progress and perspective. **International Soil and Water Conservation Research**, 11(3), 429–454, 2023.
- WEN, R.; LI, S. Spatial Decision Support Systems with Automated Machine Learning: A Review. **ISPRS International Journal of Geo-Information**, v. 12, n. 1, p. 12, 30 dez. 2022.
- XIOURAS, C., CAMELI, F., QUILLO, G. L., KAVOUSANAKIS, M. E., VLACHOS, D. G., & STEFANIDIS, G. D. Applications of Artificial Intelligence and Machine Learning Algorithms to Crystallization. **Chemical Reviews**, v. 122, n. 15, p. 13006–13042, 10 ago. 2022.
- ZHOU, L. Perspective Chapter: Big Data and Deep Learning in Hydrological Modeling. Em: **River Basin Management - Challenges and Coping Strategies [Working Title]**. [s.l.] IntechOpen, 2024.

6 CAPÍTULO 3 – INFLUÊNCIA DE VARIÁVEIS EDAFOCLIMÁTICAS NA AVALIAÇÃO ADSORTIVA DE UM BIOCARVÃO POR MODELO ENSEMBLE DE INTELIGÊNCIA ARTIFICIAL

DESTAQUES

- Investigar o mecanismo de ação do biocarvão com incorporação de variáveis ambientais e do solo.
- Estimar melhor modelo preditivo para adsorção de um agrotóxico multissítio.
- Avaliar modelo ensemble para a predição de fatores-chave de um biocarvão.
- Resultados com análises explicativas para a influência de variáveis mais influentes no processo.

RESUMO GRÁFICO



6.1 INTRODUÇÃO

Nos últimos 15 anos o desenvolvimento de pesquisas com o biocarvão (SHAHEEN et al., 2022; ZHU et al., 2023; ADENIYI et al., 2024; SONG et al., 2024; YIN et al., 2025), como alternativa ambiental tem sido intensificada. Ele é um biocomposto carbonáceo,

caracterizado na forma estável, oriundo do processo de queima de biomassa na ausência controlada de oxigênio, sob condições de temperatura programada. A otimização de sua formulação tem sido apresentada buscando-se utilização de biomassas de fontes diversificadas, calibração ótima seus parâmetros produtivos, visando minoração de sua complexidade fabril e de custos, assim como efetividade de suas funcionalidades no meio, principalmente, em remediação de poluentes (XIE et al., 2024; GOTORE et al., 2024; YAO et al., 2024), no sequestro de carbono (ZHANG et al., 2023; CHEN et al., 2024; HUANG et al., 2024), ou na regularização do solo (PALANSOORIYA et al., 2022; LV et al., 2023). Sua destinação e qualidade funcional são determinadas por atributos-chave (Gupta et al., 2023; MURTAZA et al., 2024), biomassa (ANG et al., 2023; MA et al., 2024; TANG et al., 2024), sendo influenciado ainda, por várias propriedades do solo (AZIZ et al., 2024; EBRAHIMI et al., 2024), como pH, densidade do solo, capacidade de retenção de água, fatores físico-químicos, além de fatores climáticos (XIANG et al., 2023; LI et al., 2024), incluindo aí, temperatura e umidade relativa.

Aplicação da modelagem computacional do biocarvão com ferramentas de inteligência artificial tem sido empregada com sucesso para investigar o biocarvão (GHANBARI et al., 2023; REX et al., 2023; ABHAYASIMHA et al., 2024; SU; JIANG, 2024; XIE et al., 2025), principalmente para aprimorar os métodos produtivos, funcionalidades e impactos ambientais. A modelagem de acervos de experimentos de bancada pode ser ampliada, investigando-se adições de novos parâmetros, variáveis, ou condições ambientais ampliadas, que possam influenciar nas etapas ou atributos produtivos, como biomassa, tempo de pirólise, temperatura de queima, dentre outros.

Recentemente, com a utilização de modelos de inteligência artificial com aprendizado profundo (DL) em hierarquias ensemble (ZHENG et al., 2022; GABSI et al., 2023; GOU et al., 2025; MITTAL et al., 2025; YAQIB; LEE, 2023), a identificação e correlacionamento entre fatores críticos que afetam o desempenho do biocarvão podem ser personalizados, visando otimização de seus resultados. Lacunas ainda pouco investigadas, como a afetação do biocomposto no solo, em função do tempo (*aging*) e variação de suas condições, podem ser analisadas, descobrindo-se novos padrões, a partir de simulações com inclusões de novas variáveis e apresentação de grandes conjuntos de dados, aprimorando as propriedades físicas e químicas do biocarvão, direcionando-o para destinações ambientais e agrícolas específicas. Agora, a introdução de variáveis climáticas e de solo na sua produção representa uma nova intersecção inovadora e vantajosa, permitindo a análise de interações complexas entre as diversas variáveis que afetam seu

rendimento e qualidade, pouco analisadas em laboratório, pelas condições complexas de manipulação e avaliação. Questões cruciais, como sua viabilidade econômica e ampliação do espectro de atuação, poderão ser mais bem avaliadas.

Este estudo propõe avaliar a capacidade adsortiva de um biocarvão, com a introdução de variáveis intrínsecas, como a condutividade hidráulica do solo e a evapotranspiração real diária, na geração das suas curvas isotermas, através da modelagem com algoritmos de aprendizado profundo, comparativamente aos resultados experimentais. A viabilidade dessa pesquisa propõe estimular a ampliação de pesquisas nessa direção, amparadas por poderosas ferramentas de IA.

O artigo está estruturado em quatro etapas. A segunda etapa, de Materiais e Métodos, detalha a sequência de desenvolvimento da pesquisa, os procedimentos de geração de dados, descrição da região de estudo, construções dos modelos de inteligência artificial (IA), seleção de melhores hiperparâmetros, desempenho dos modelos avaliados. A etapa de Resultados e Discussões compreende breve análise do conjunto de dados originais e sintéticos, avaliação da predição pelo melhor modelo de IA, análises explicativas dos resultados obtidos, buscados por esta pesquisa. A etapa de Conclusões resume o estudo e oferece recomendações para estudos futuros.

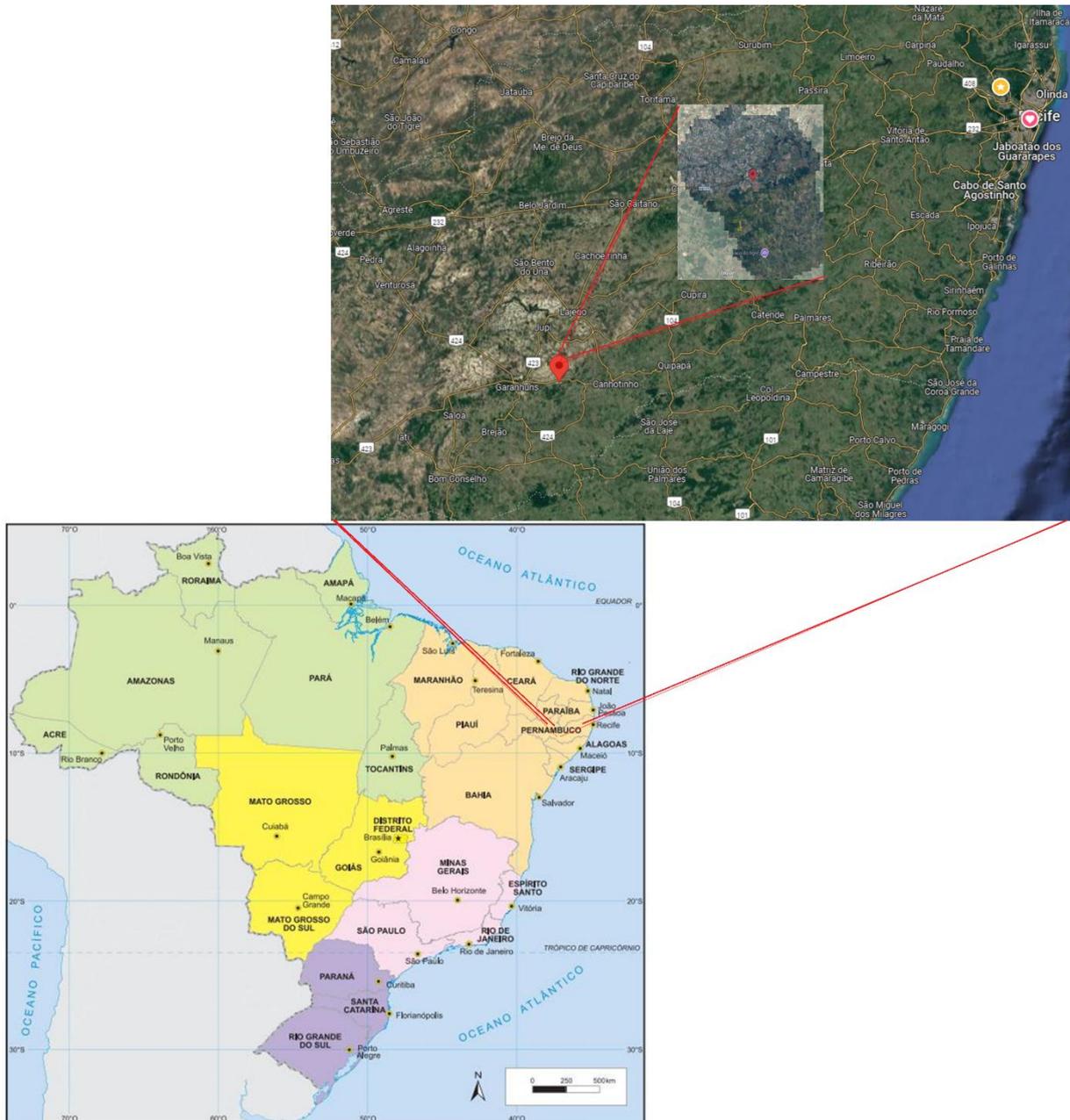
6.2 MATERIAIS E MÉTODOS

6.2.1 Área de estudo

Para este estudo, os dados de condutividade hidráulica saturada – K_s (mm/s) e evapotranspiração real diária – ET_r (mm/dia) foram obtidos de dois experimentos realizados em uma localidade de origem agrícola, com aproximadamente 2.500 m², comumente utilizada para pastagem de gado, onde a vegetação cultivada é a *Braquiária decumbens* Stapf. A área fica localizada na fazenda Riacho do Papagaio, município de São João (Figura 1), na mesorregião do Agreste Meridional, no estado de Pernambuco, de coordenadas geográficas 8° 52' 30" S, e 36° 22' 00" W, em altitude de 705 m. O clima observado é predominantemente o As', sendo equivalente a clima quente e úmido, conforme classificação de Köeppen (Beck et al., 2018). O período chuvoso é previsto na fase de outono/inverno, com início em dezembro/janeiro e término até setembro. A precipitação

média anual registrada na literatura para a região é de aproximadamente 1.310 mm. Por falta de amostras experimentais desta área, os valores de concentração do soluto após equilíbrio na fase líquida - C_{eq} (mg/l) e capacidade de adsorção - S (mg/kg), utilizados neste estudo, essenciais para a produção e avaliação adsortiva do biocarvão, são oriundos de ensaios experimentais realizados na Usina São José Agroindustrial, localizada na zona rural do município de Igarassu, estado de Pernambuco, destinado para a produção agrícola de cana-de-açúcar, nas coordenadas geográficas $7^{\circ} 47' 59,02''$ S e $35^{\circ} 0' 18,45''$ W. O solo da área é caracterizado pela predominância de Argissolo Amarelo distrófico (Aad). O clima é o tropical chuvoso, com verão seco. O período chuvoso ocorre na janela de fevereiro a outubro, com precipitação média anual de 1.634 mm, com vegetação predominante do tipo floresta Subperenifólia, com trechos em Subcaducifólia e cerrado/floresta.

Figura 1: Localização geográfica da área de estudo (coleta de ETr e Ks), em São João/PE.



Fonte: Autor (2024)

6.2.2 Geração de dados e delineamento experimental

Os dados experimentais de ETr e NDVI utilizados nesta pesquisa para valores estimados de ETr foram gerados pelo sistema MOD16A2, da NASA (MOD16A2, 2024). Ele forneceu estimativas de evapotranspiração derivadas do modelo de balanço da vegetação da área de estudo, com resolução espacial de 500 metros e intervalo temporal de 8 dias, no período de 2023, com imageamento processado pela plataforma Google Earth Engine (GEE, 2024). Em seguida, foi procedido ao delineamento experimental estruturado e recorte para a área de estudos, utilizando-se a ferramenta de modelagem de imagens QGIS®

(QGIS, 2024). Os dados geoprocessados foram tabulados para listas temporais em duas planilhas, com 1.055 amostras cada, onde ETr continha duas colunas, sendo uma com pontos de vetor e outra com pontos de ETr, e NDVI contendo uma coluna com pontos de NDVI.

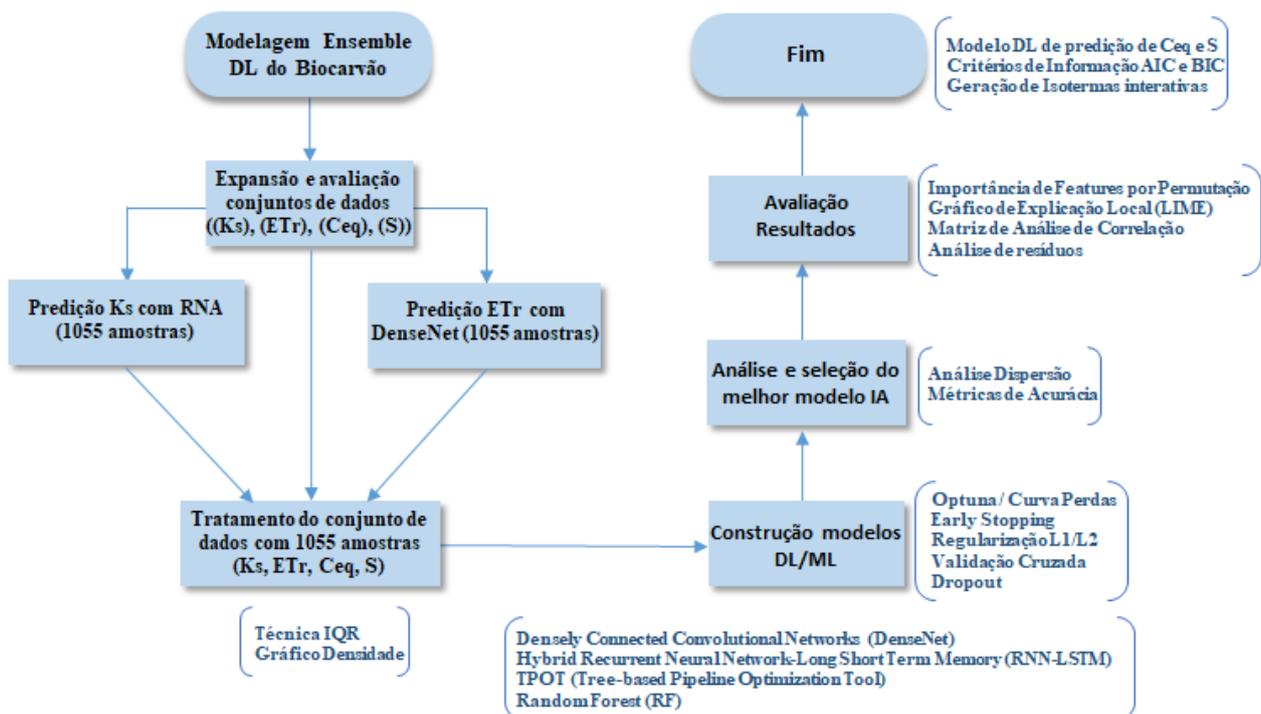
Os dados experimentais de condutividade hidráulica saturada (K_s) do solo são oriundos de experimento realizado na área de estudo, com um pequeno conjunto amostral (areia, silte, argila, densidade do solo), expandidos para 1.055 amostras sintéticas, visando estimar K_s .

E por fim, os dados experimentais de C_{eq} (mg/l) e S (mg/kg) foram obtidos de um biocarvão produzido pela queima lenta a 700°C , com biomassa oriunda do bagaço de cana-de-açúcar, extraído na Usina São José Agroindustrial. O biocarvão foi avaliado quanto sua capacidade de adsorver o herbicida Imazapic[®], em um solo de plantio de cana-de-açúcar, através da geração de curvas isotermas. O experimento produtivo do biocarvão gerou, dentre outras variáveis, 257 medições de concentração do soluto, após o equilíbrio na fase líquida (C_{eq}), e de capacidade de adsorção (S), essenciais para avaliar a capacidade adsorptiva de um biocarvão sobre solo contaminado com pesticidas. Estes dados foram então expandidos para 1.055 amostras sintéticas, visando equilibrar a quantidade de amostras por variáveis.

6.2.3 Planejamento do desenvolvimento da pesquisa

A descrição da metodologia, etapas e métodos aplicados neste estudo é apresentada resumidamente na Figura 2, abaixo, e detalhada nas seções seguintes.

Figura 2: Esquema metodológico de desenvolvimento da pesquisa.

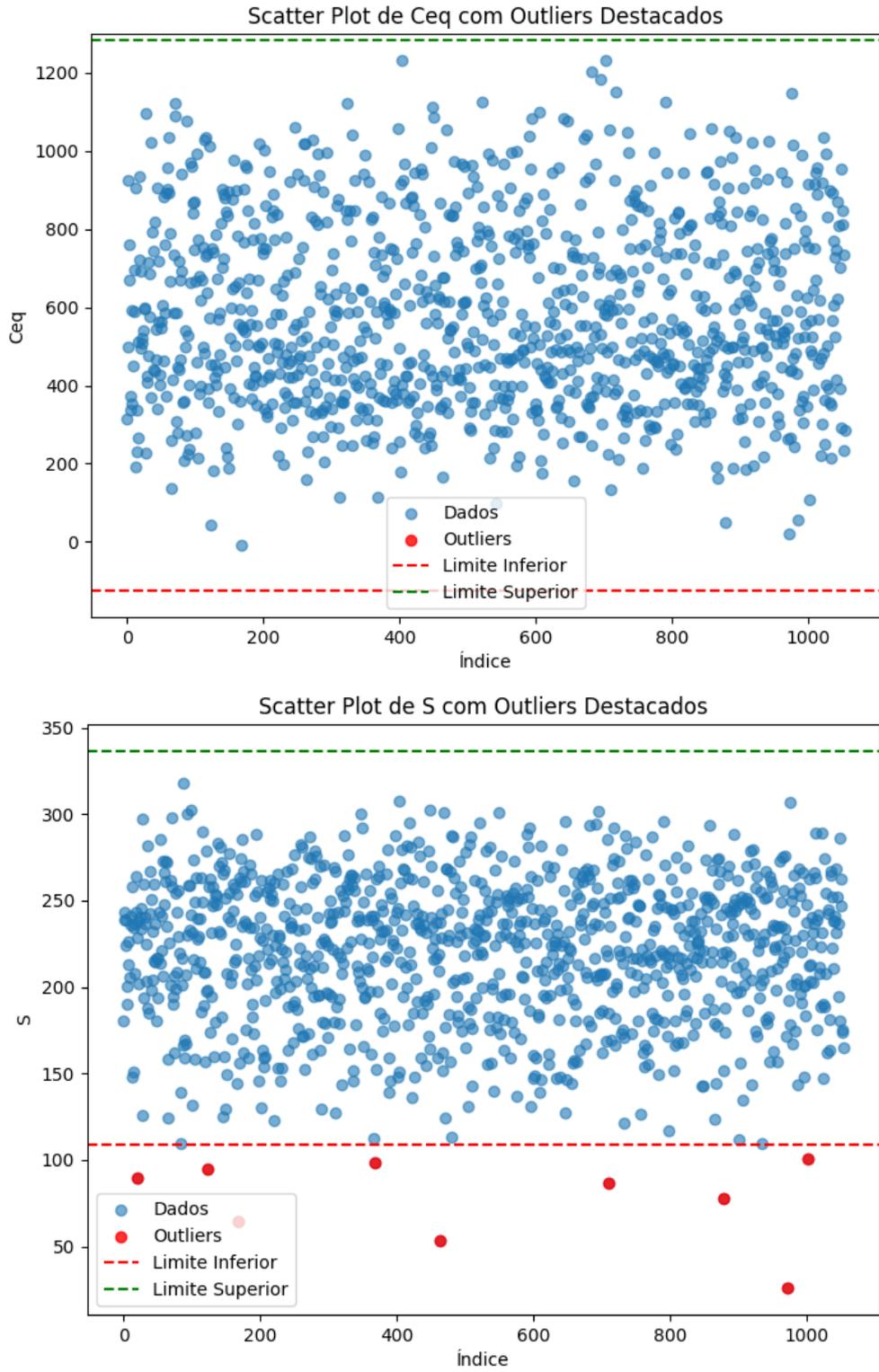


Fonte: Autor (2024)

6.2.4 Avaliação das variáveis de estudo

O conjunto de dados integrados (Ks, ETr, Ceq, S), cada um contendo quatro colunas de 1.055 amostras, foi submetido à análise de pontos extremos (*outliers*), utilizando a técnica de interquartil – IQR (KORKMAZ, 2023), permitindo visualizar 7 pontos nos valores de ETr, 4 pontos nos valores de Ks, 8 pontos nos valores de Ceq (Figura 3a) e nenhum ponto extremo em S (Figura 3b), totalizando 19 pontos extremos no conjunto amostral.

Figura 3: Distribuição espacial dos dados sintéticos – a) variável Ceq; b) variável S.



Fonte: Autor (2024)

6.2.5 Construção do modelo Ensemble

Uma arquitetura hierarquizada, composta por um ensemble com três modelos neurais, sendo uma rede neural artificial (RNA) e duas redes neurais profundas (DL) DenseNet, foi desenvolvida para estimar parâmetros-chave das condições adsorptivas de um experimento de biocarvão, com adição das variáveis de evapotranspiração e condutividade hidráulica saturada do solo.

Uma rede neural artificial (RNA), previamente selecionada, foi treinada para estimar valores da condutividade hidráulica saturada (K_s). Em paralelo, uma DenseNet, também antecipadamente avaliada, foi implementada para estimar a evapotranspiração real diária (ET_r).

Os valores ET_r e K_s estimados foram integrados aos dados sintéticos de C_{eq} e S , alimentando uma terceira rede, com melhores métricas de acurácia, selecionada entre quatro modelos de IA testados. Este modelo de aprendizado profundo (DL) analisou a relação e influência dessas variáveis preditoras (K_s , ET_r, C_{eq} e S) sobre as estimações das variáveis-alvo C_{eq} e S . Isotermas de Freundlich e Langmuir foram então geradas pelo ensemble, visando avaliar a capacidade adsorptiva do biocarvão, comparativamente às curvas de isotermas originais experimentais.

6.3 RESULTADOS E DISCUSSÕES

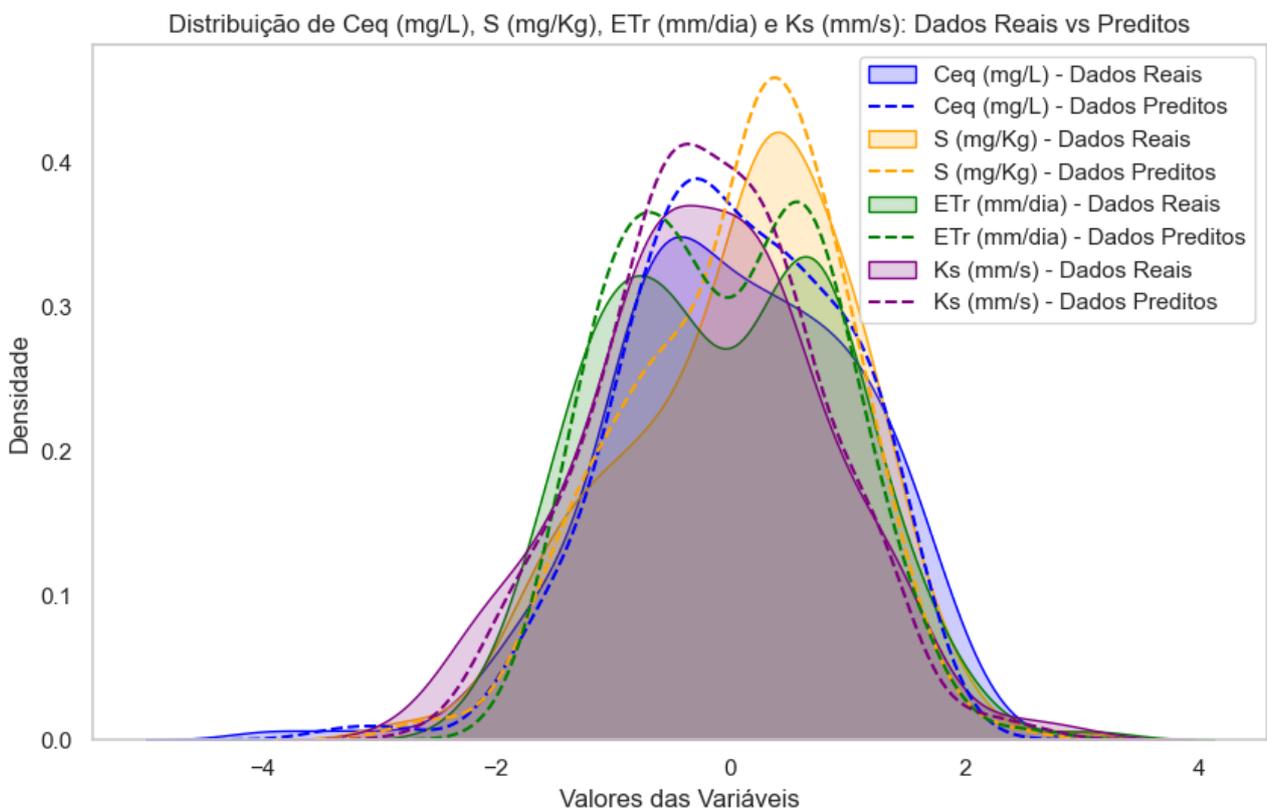
6.3.1 Análise do conjunto amostral de dados

A manutenção ou poda dos pontos extremos foi procedida através da avaliação da análise exploratória de dados, onde a manutenção indicou menor impacto na acurácia, onde alguns pontos extremos apresentavam importância na representação da variabilidade, ante o conjunto de dados. A pequena diferença entre valores reais e sintéticos, indicada pela média das variáveis, foi razoavelmente mantida pela expansão dos dados. O intervalo de distribuição para os valores sintéticos mostrou-se ligeiramente expandido em relação aos originais, sugerindo novo tratamento pós-processamento dos dados sintéticos, caso necessário, mas plenamente viáveis para a realização deste estudo.

O desvio-padrão nos dados sintéticos é um pouco menor que os originais, indicando menos dispersão. As medianas dos dados sintéticos são muito próximas das amostras originais, reforçando a capacidade de preservação da distribuição central dos dados. Os testes de Kolmogorov-Smirnov sugeriram pequena distorção na distribuição, mas viáveis para a modelagem.

A análise da distribuição de densidade entre as variáveis originais e preditas demonstrou o impacto entre variáveis preditoras e alvo, modeladas pela DenseNet, onde a C_{eq} apresentou boa sobreposição, indicando boa captura da distribuição. Já a distribuição mostrou que a distribuição foi razoável, mas com consistência para a estimação. A análise dos valores de E_{Tr} sugere um pequeno viés, mas sem prejudicar sua utilização. As curvas real e predita de K_s apresentam desvio mais acentuado, tendendo a maior esforço na sua captura. No geral, as aproximações Gaussianas indicaram distribuições normais, coerentes com o comportamento de variáveis ambientais contínuas. O resultado indica que o modelo não representou com excelência a relação K_s/E_{Tr} com as variáveis-alvo C_{eq} e S , causadas por necessidade de melhorias no modelo, treinamento ou quantidade de dados ou ajustes na arquitetura do modelo.

Figura 4: Densidade espacial dos dados sintéticos.



6.3.2 Desempenho do modelo Ensemble

O modelo de IA escolhido para modelar os valores preditores (Ceq, S, Ks e ETr) para estimação das variáveis-alvo foi selecionado da avaliação de quatro algoritmos de aprendizado, utilizando os resultados de suas métricas de acurácia para tal. Foi considerado como premissa para investigar esses modelos, pesquisas na literatura e características dos dados. Os modelos TPOT (WEN; LI, 2022), Redes Neurais Artificiais (MOZAFFARI et al., 2024), Random Forest (KANG et al., 2023) e DenseNet (SINGH et al., 2023) utilizaram-se de arquiteturas similares, onde o mesmo conjunto de dados (Ceq, S, ETr e Ks), dividido em treinamento (80%) e testes (20%) aplicaram a biblioteca Optuna, que definiu seus principais hiperparâmetros, além de técnicas como Dropout, BatchNormalization, validação cruzada, EarlyStopping, regularização L1/L2, para gerar as métricas de acurácia, gráficos e análises explicativas para estimação das variáveis-alvo.

Os quatro modelos apresentaram resultados próximos (Tabela 1), onde a DenseNet foi escolhida.

Tabela 1: Resultados das métricas de acurácia para os quatro modelos de Inteligência Artificial avaliados.

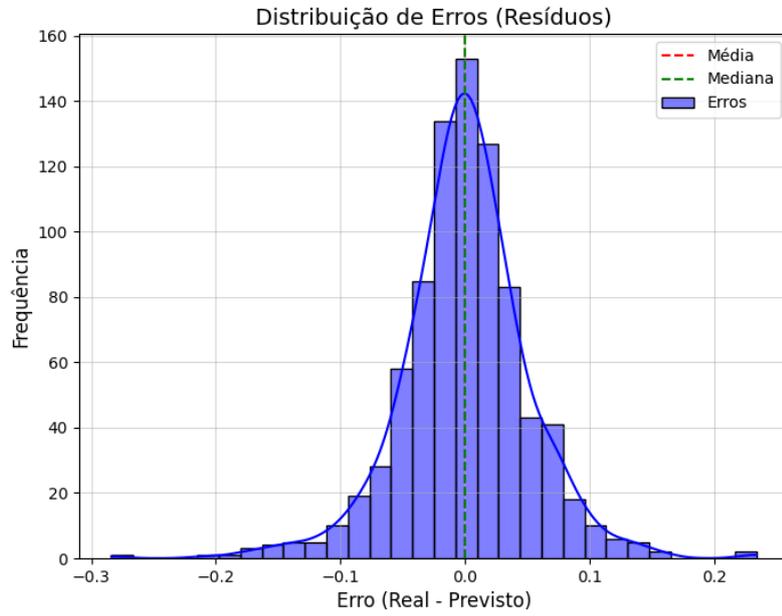
	Método de Busca	MSE	RMSE	MAE	R ²	CC
DenseNet	Optuna	0,02	0,13	0,09	0,99	1,00
RNN-LSTM	Optuna	0,02	0,15	0,11	0,98	1,00
RF	Optuna	0,03	0,18	0,13	0,97	0,99
TPOT	Optuna	0,03	0,18	0,14	0,97	0,99

Fonte: Autor (2024)

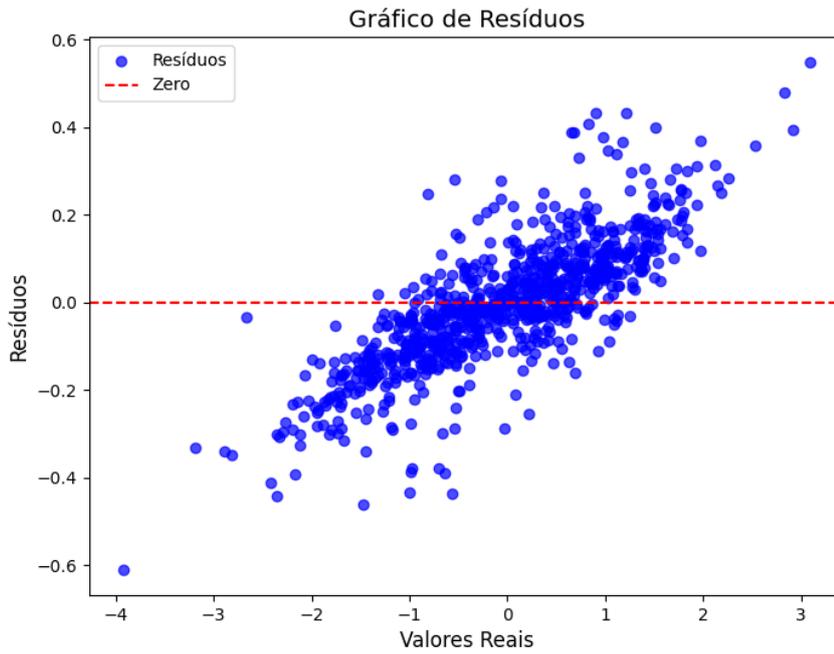
Os erros resultantes da modelagem foram avaliados, onde a qualidade preditiva e presença de viés foram analisadas (Figura 5). O histograma (a) do modelo indicou que a distribuição dos erros foi normal e simétrica, com centro ao zero, com ausência de viés significativo e boa generalização. A presença de assimetria severa não foi identificada, pela proximidade entre a média e a mediana. O gráfico de resíduos (b) mostrou falta de padrão evidente nos erros, indicando não existência de variância constante dos erros (distribuição em torno do eixo $y=0$), detectando ocorrência de erros nos valores extremos, sem prejudicar a modelagem, com inclusão dos valores de ETr e Ks.

Figura 5 – Gráficos dos dados sintéticos – a) resíduos; b) dispersão.

(a)



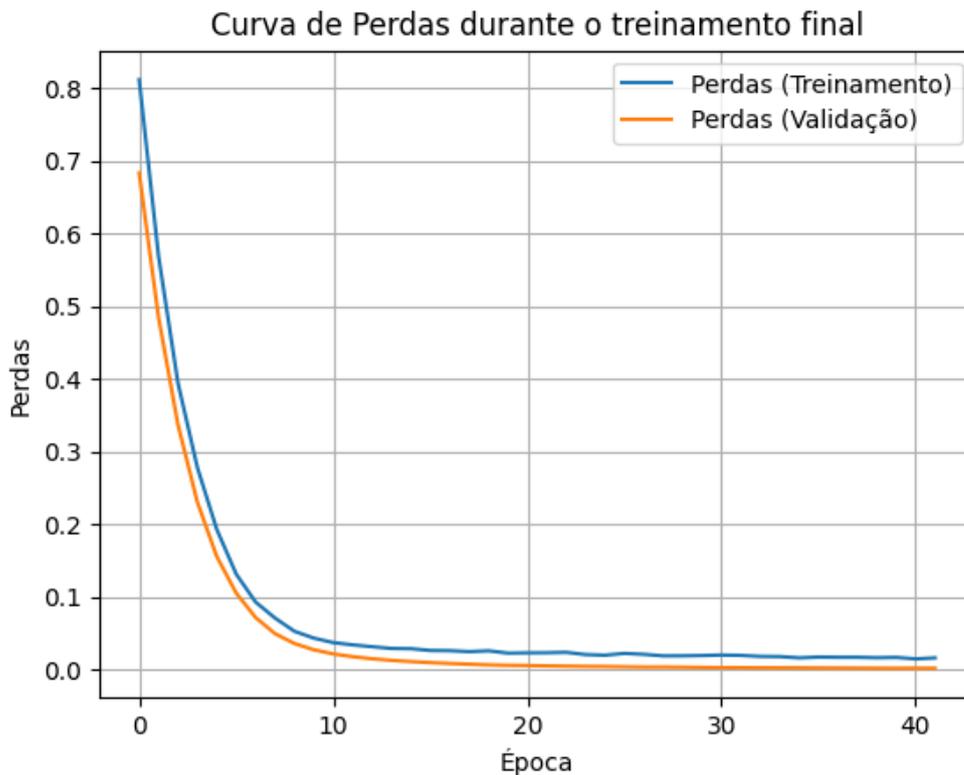
(b)



Fonte: Autor (2024)

A curva de perdas (Figura 6) corroborou os resultados de resíduos e dispersão de erros, na fase de treinamento e validação, indicando rápida convergência e aprendizado eficiente, estabilizando as perdas, indicando que o modelo atingiu ótimo aprendizado, sem a presença prejudicial de sobreajuste, confirmando que a inclusão de ETr e Ks não prejudicou o aprendizado, sem presença excessiva de ruídos.

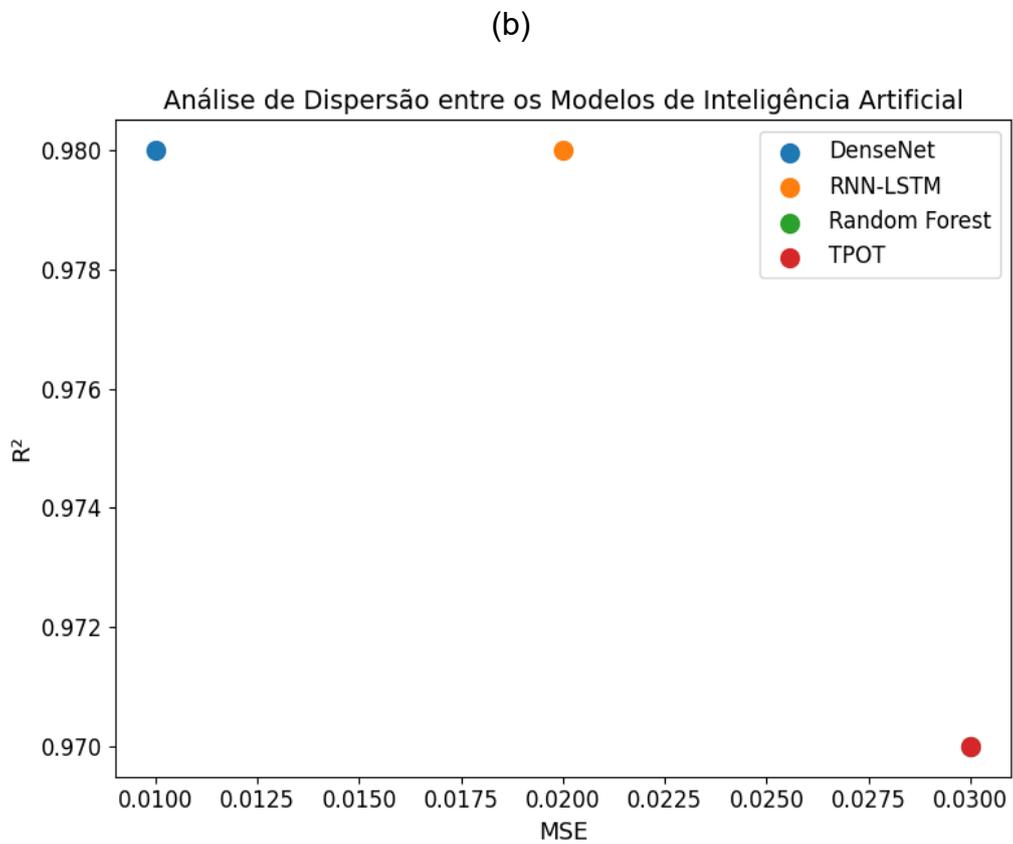
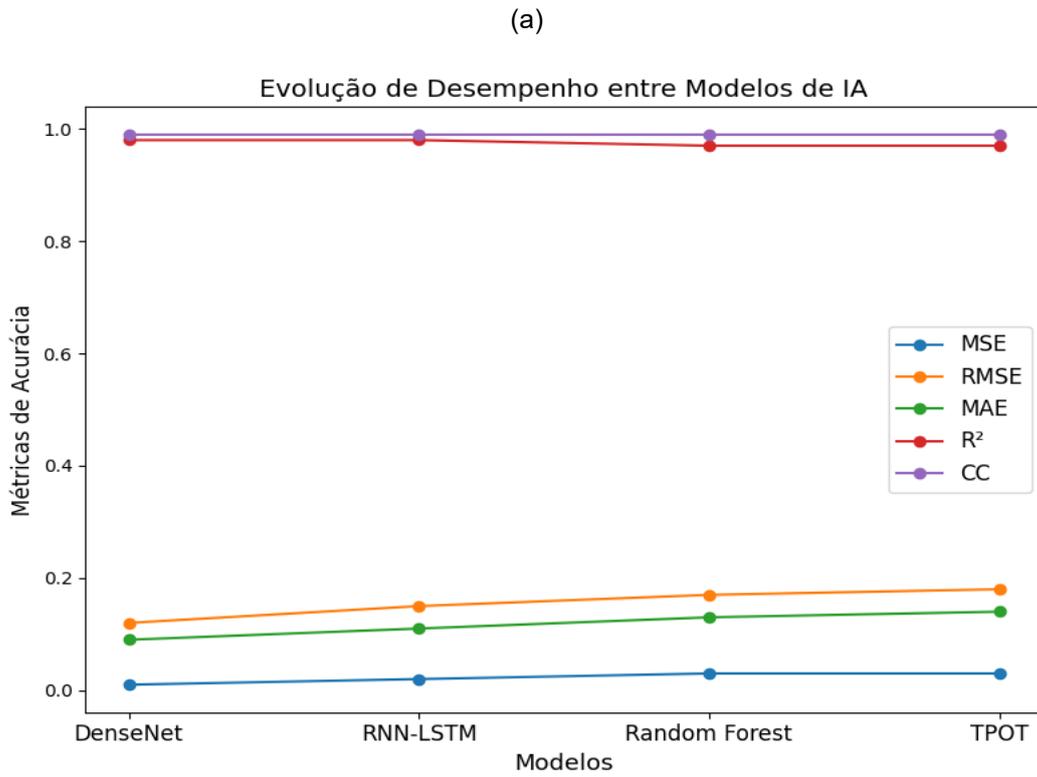
Figura 6 – Gráfico de Curvas de Perdas durante o treinamento do modelo DenseNet.



Fonte: Autor (2024)

A geração dos gráficos (Figura 7) de desempenho geral (a) confirmou o desempenho superior da DenseNet, apresentando os melhores valores de CC e R^2 (b), indicando boa generalização, com valores crescentes. TPOT e RF apresentaram desempenho inferior a DenseNet, principalmente pela limitação em capturar as relações não lineares complexas entre ETr e Ks.

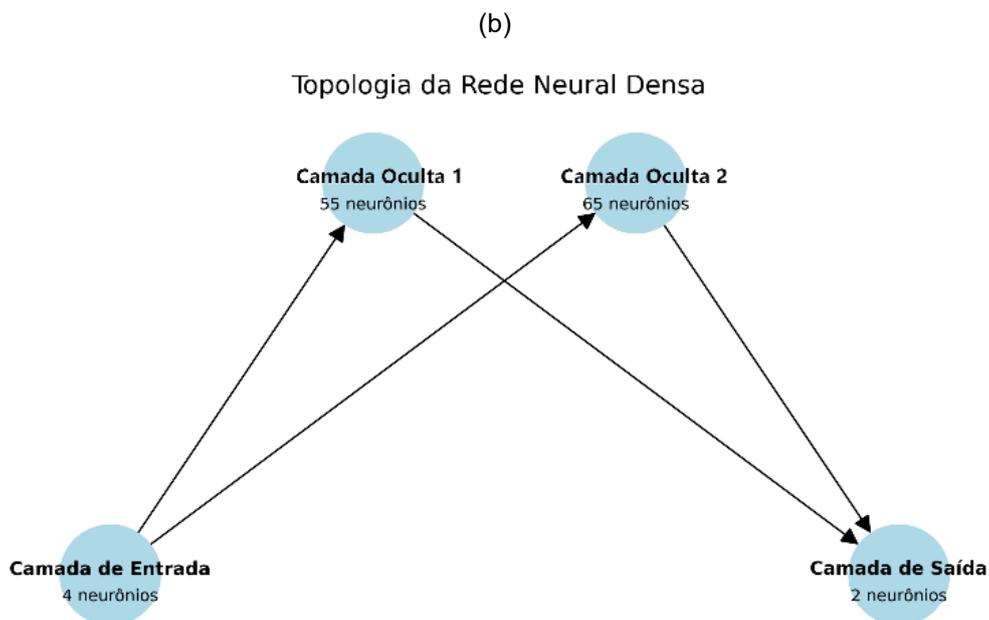
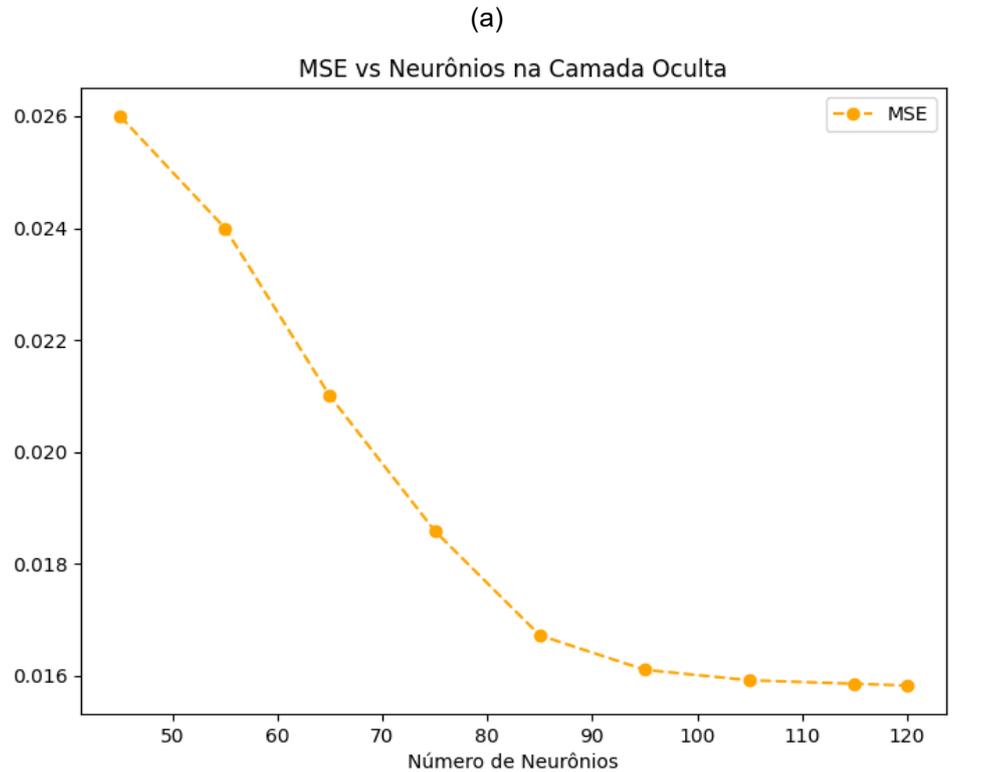
Figura 7 – Gráficos dos modelos IA avaliados – a) desempenho; b) acurácia MSE vs R².



Fonte: Autor (2024)

A arquitetura da DenseNet foi analisada (Figura 8) na evolução de sua capacidade preditiva, onde o platô de convergência por volta de cem neurônios (a), indicando que a parada poderia ter sido antecipada, mas bem calibrada. A relação direta entre os nós da rede (b) confirma as relações diretas entre as variáveis predictoras e alvo.

Figura 8 – Gráficos dos modelos IA avaliados – a) acurácia MSE vs Neurônios; b) topologia.

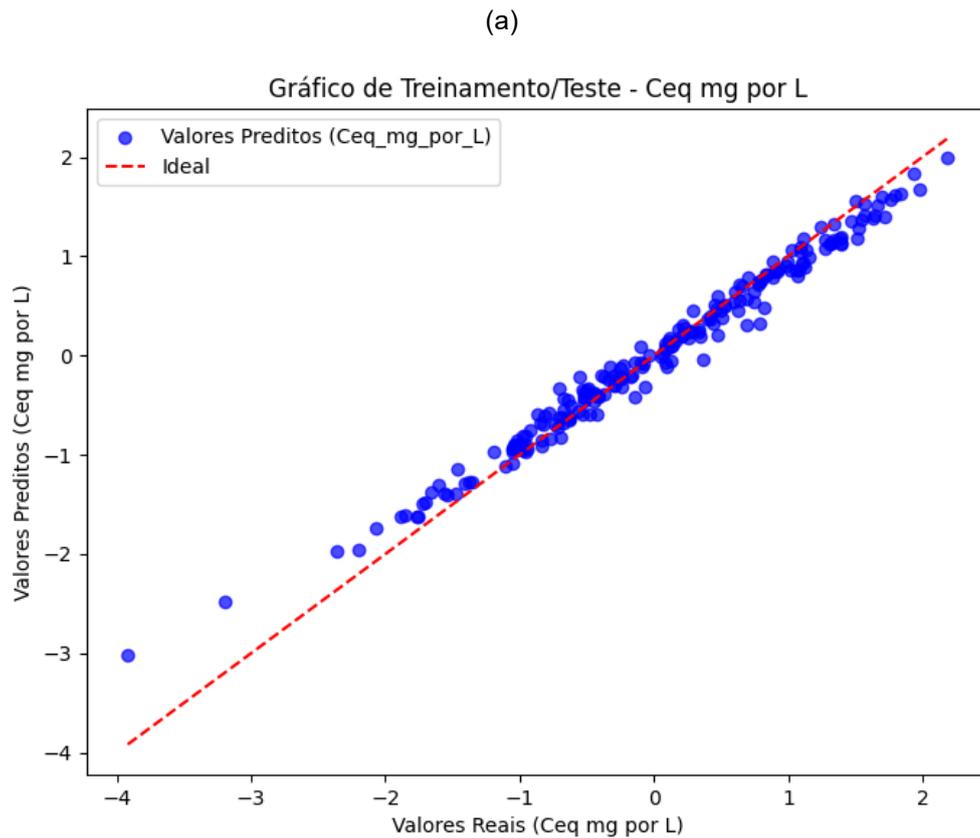


Fonte: Autor (2024)

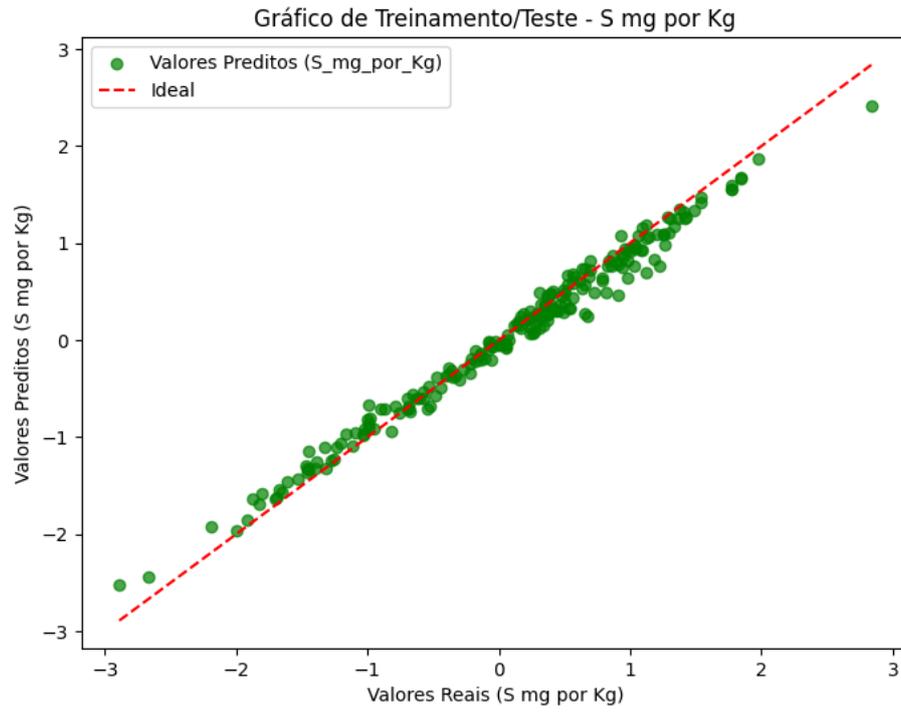
A aderência dos valores preditos de C_{eq} , S , E_{Tr} e K_s (Figura 9) foi avaliada individualmente nas fases de treinamento e testes, indicando que a DenseNet capturou corretamente as relações entre todas as variáveis, com baixa dispersão e sem grandes erros sistemáticos.

Os valores de C_{eq} (a) foram bem modelados, de boa taxa de confiabilidade. A variável S (b) apresentou boa aderência, com uma pequena dispersão, mas sem desvios graves. O gráfico da variável E_{Tr} (c) mostrou que ocorreu dispersão minorada para valores negativos, ressaltando dificuldade em faixas específicas, mas sem afetar a estimação, de forma consistente. A estimação mostrada de K_s (d) indicou boa captura na modelagem, sem grandes distorções.

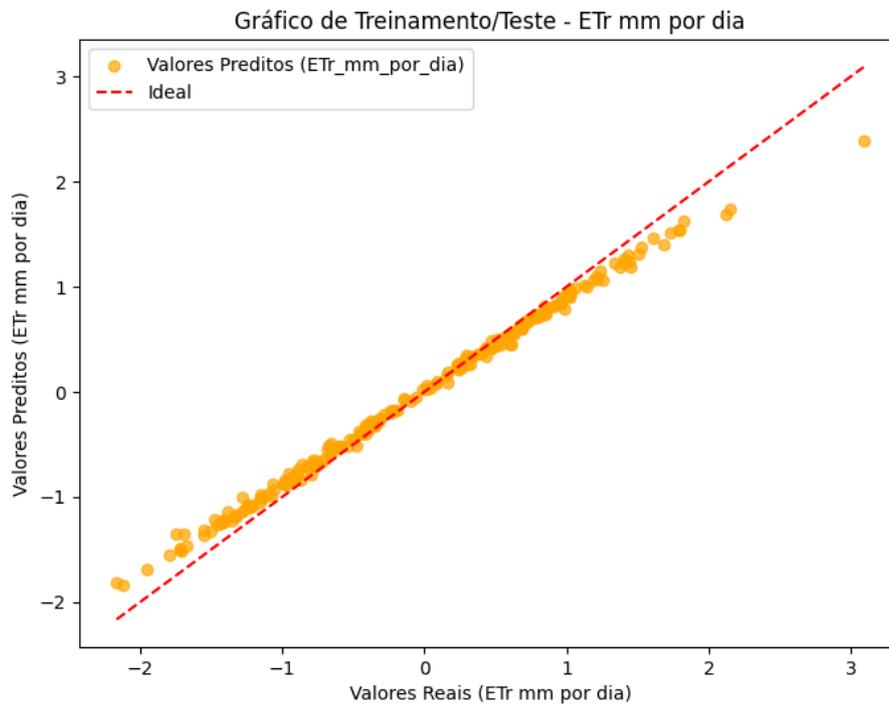
Figura 9 – Desempenho da DenseNet nas fases de treinamento e teste – a) C_{eq} ; b) S ; c) E_{Tr} ; d) K_s .



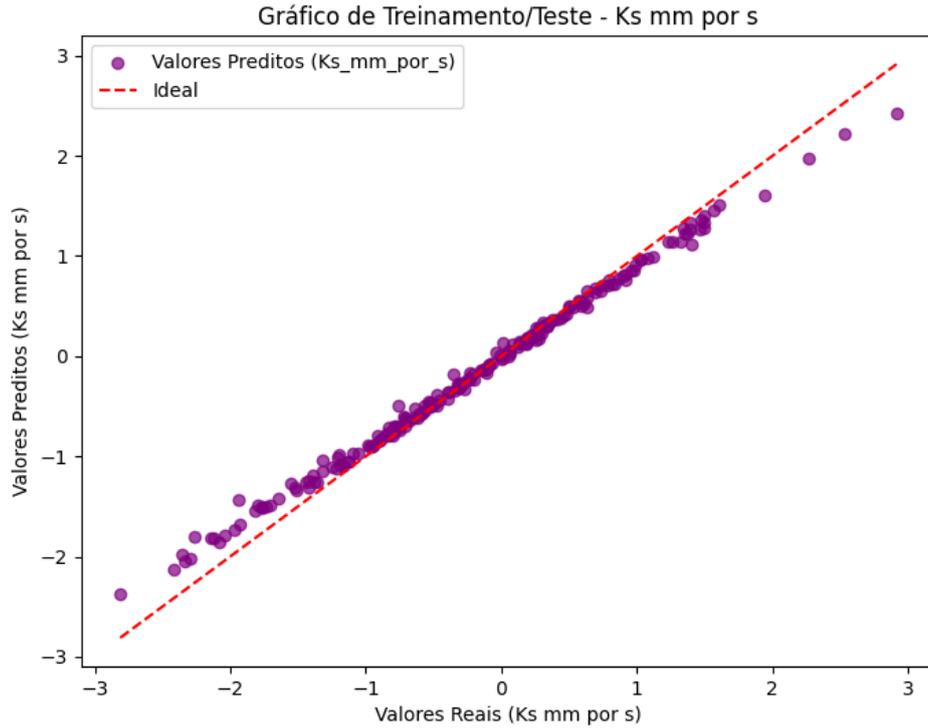
(b)



(c)



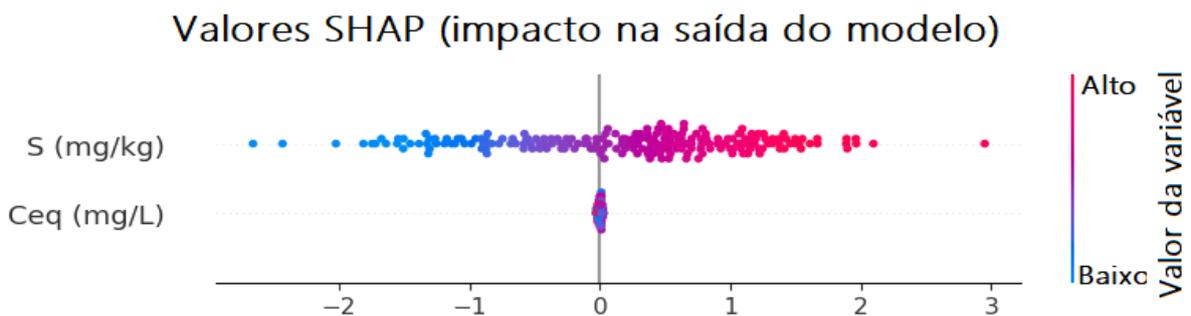
(d)



Fonte: Autor (2024)

A análise SHAP (Figura 10) avaliou especificamente o impacto das variáveis C_{eq} e S na saída no modelo DenseNet. O gráfico indicou que S contribuiu notadamente para a estimação, com variações. A variável C_{eq} apresentou distribuição altamente centralizada, impacto menor, mas estável, ressaltando seu papel na definição das curvas isotermas.

Figura 10 – Análise SHAP (direção, complexidade e relacionamento) entre C_{eq} e S estimadas.

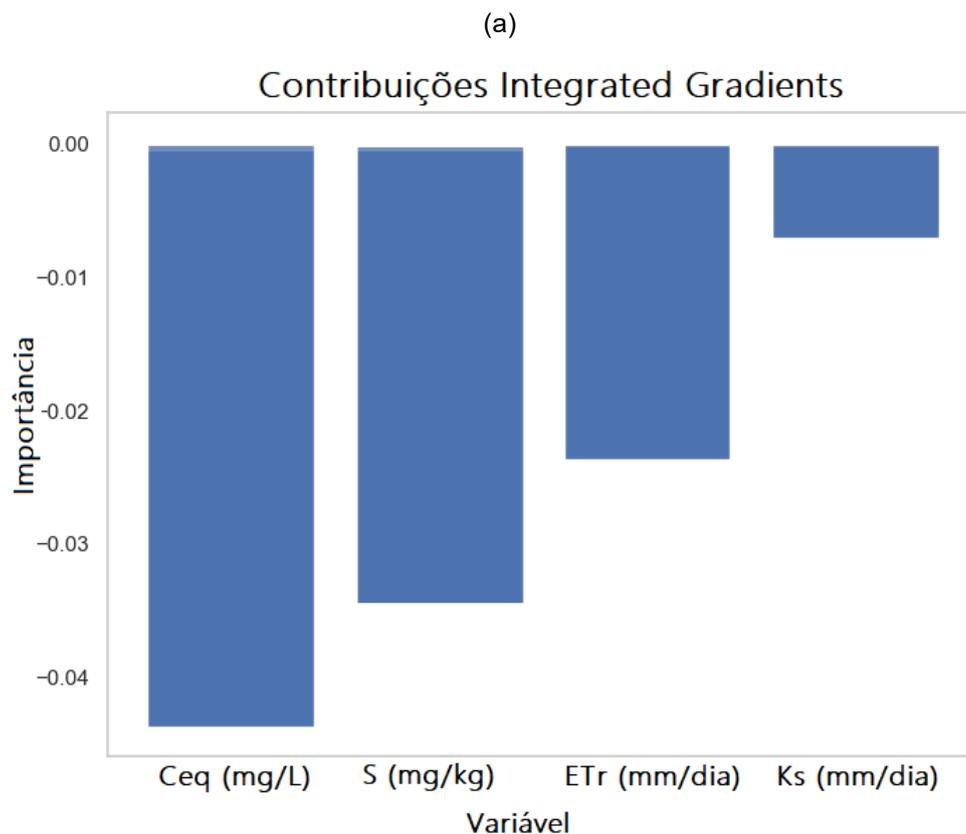


Fonte: Autor (2024)

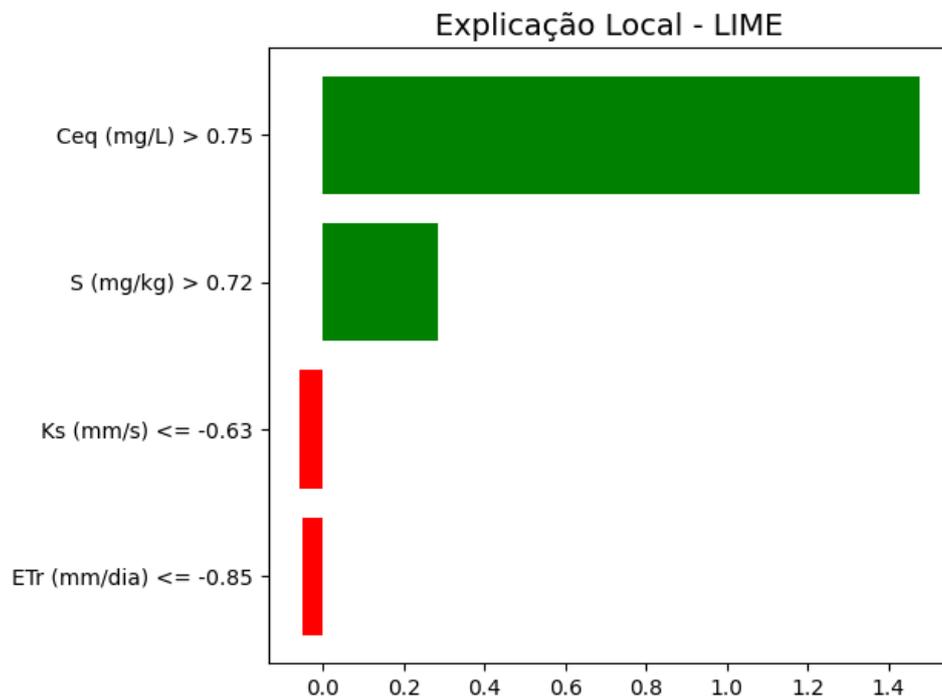
A influência das variáveis foi analisada através das técnicas de Contribuições (a) e Explicação Local (b) comparativamente (Figura 11). Os maiores impactos na estimação foram confirmados pelos valores negativos de C_{eq} e S , corroborando o sentido e importância relativa observada no gráfico SHAP.

O gráfico de Integrated Gradient fornece uma visão global do modelo, e indicou que C_{eq} e S influenciaram mais fortemente. Isto foi corroborado pelo gráfico de Explicação Local (LIME), através da análise localizada de cada variável na estimativa, onde C_{eq} e S tiveram maiores influências na estimação, embora tenham confirmado as contribuições de E_{Tr} e K_s , representados pelos valores negativos.

Figura 11 – Análises explicativas da estimação – a) Contribuições Integrated Gradients; b) LIME.



(b)

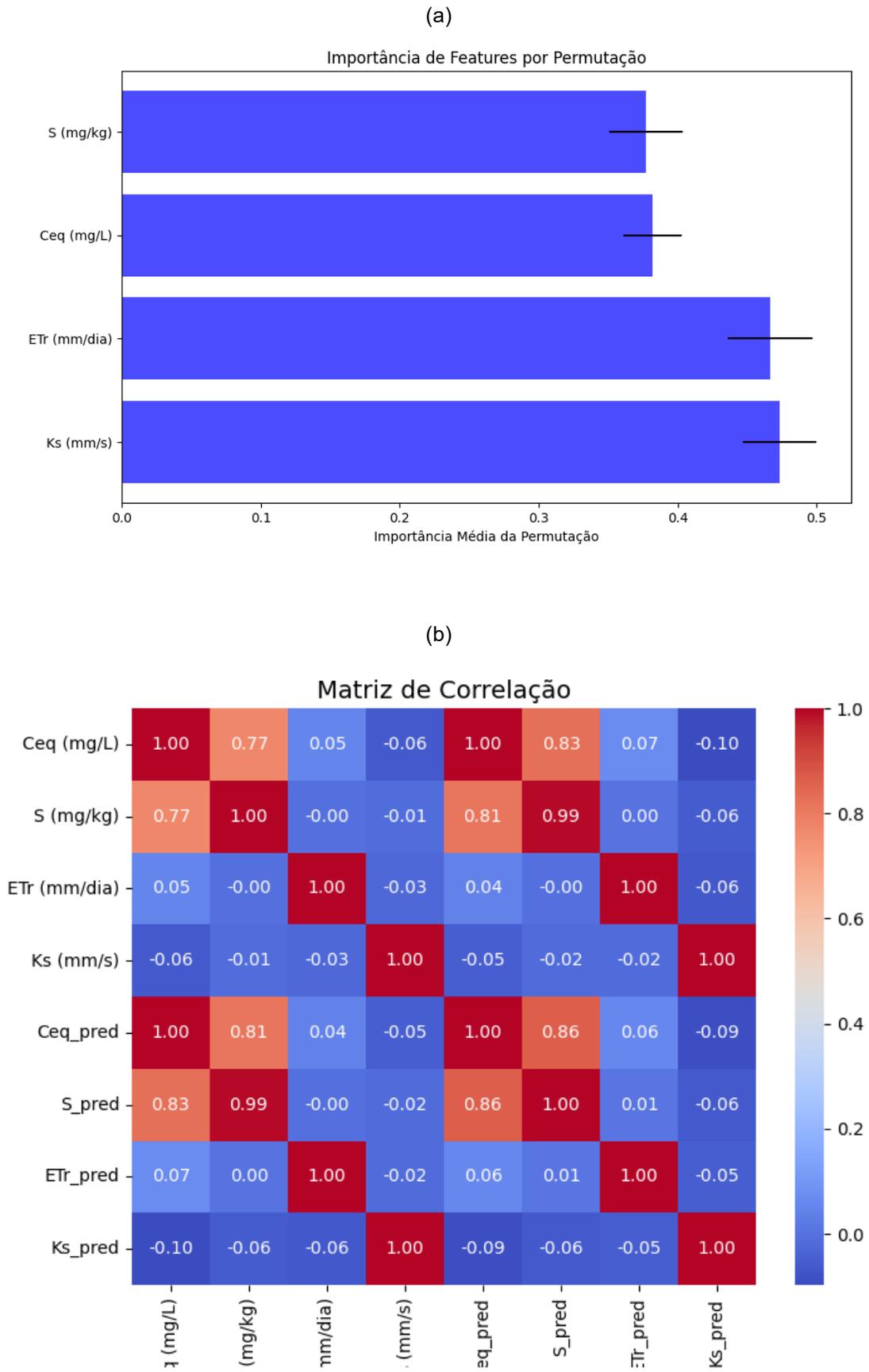


Fonte: Autor (2024)

Em outra análise realizada para investigar a correlação e influência das variáveis na predição do model DenseNet, foi realizada com a geração dos gráficos de Importância das variáveis via Permutação (Figura 12a) e Matriz de Correlação (Figura 12b). A Importância das Features por Permutação confirma que Ceq e S dominam o processo de adsorção e modelagem das isotermas, onde ETr e Ks contribuíram para uma boa aproximação. A Matriz de Correlação confirma o relacionamento intrínseco entre Ceq e S ($r = 0,77$), alta correlação entre Ceq estimado e Ceq original ($r = 0,98$), S original e S estimado ($r = 0,93$), fraca correlação entre ETr e as variáveis-alvo, assim como Ks, que mostrou uma associação levemente negativa, de menor impacto.

O resultado geral indicou que a DenseNet conseguiu refletir as relações entre as variáveis, confirmando as influências diferenciadas entre cada variável na estimativa do modelo.

Figura 12 – Análises explicativas da estimação – a) Importância por permutação; b) Matriz de correlação.



Fonte: Autor (2024)

6.3.3 Avaliação da capacidade adsorptiva pelas Isotermas Freundlich e Langmuir

A análise comparativa entre os valores experimentais de C_{eq} e S foi utilizada pela DenseNet para calcular a cinética de adsorção com os valores estimados de C_{eq} e S (Figuras 13a e 13b), comparativamente às realizadas por um biocarvão (Figuras 13c e 13d), oriunda do experimento com o herbicida Imazapic.

Os valores das isotermas estimadas de Freundlich (13a) e Langmuir (13b) demonstraram boa correspondência com os valores preditos, validando a eficácia da modelagem com o aprendizado profundo, com a investigação da adição das variáveis de E_{Tr} e K_s . As curvas estimadas (13d) refletiram o comportamento das curvas experimentais (13c), onde Freundlich apresentou um ajuste levemente superior nos pontos intermediários.

A comparação explicativa (Tabela 2) entre as isotermas de Freundlich e Langmuir foi realizada através da aplicação dos critérios AIC - Akaike Information Criterion (OBEY et al., 2022) e BIC - Bayesian Information Criterion (LI et al., 2024), procurando refletir de forma consistente o comportamento físico-químico da adsorção. O critério combinado AIC/BIC gerou valores penalizados e com ajuste matemático, a partir das estimações de C_{eq} e S pelo modelo neural profundo, onde valores menores indicam melhor ajuste, penalizando sistemas mais complexos.

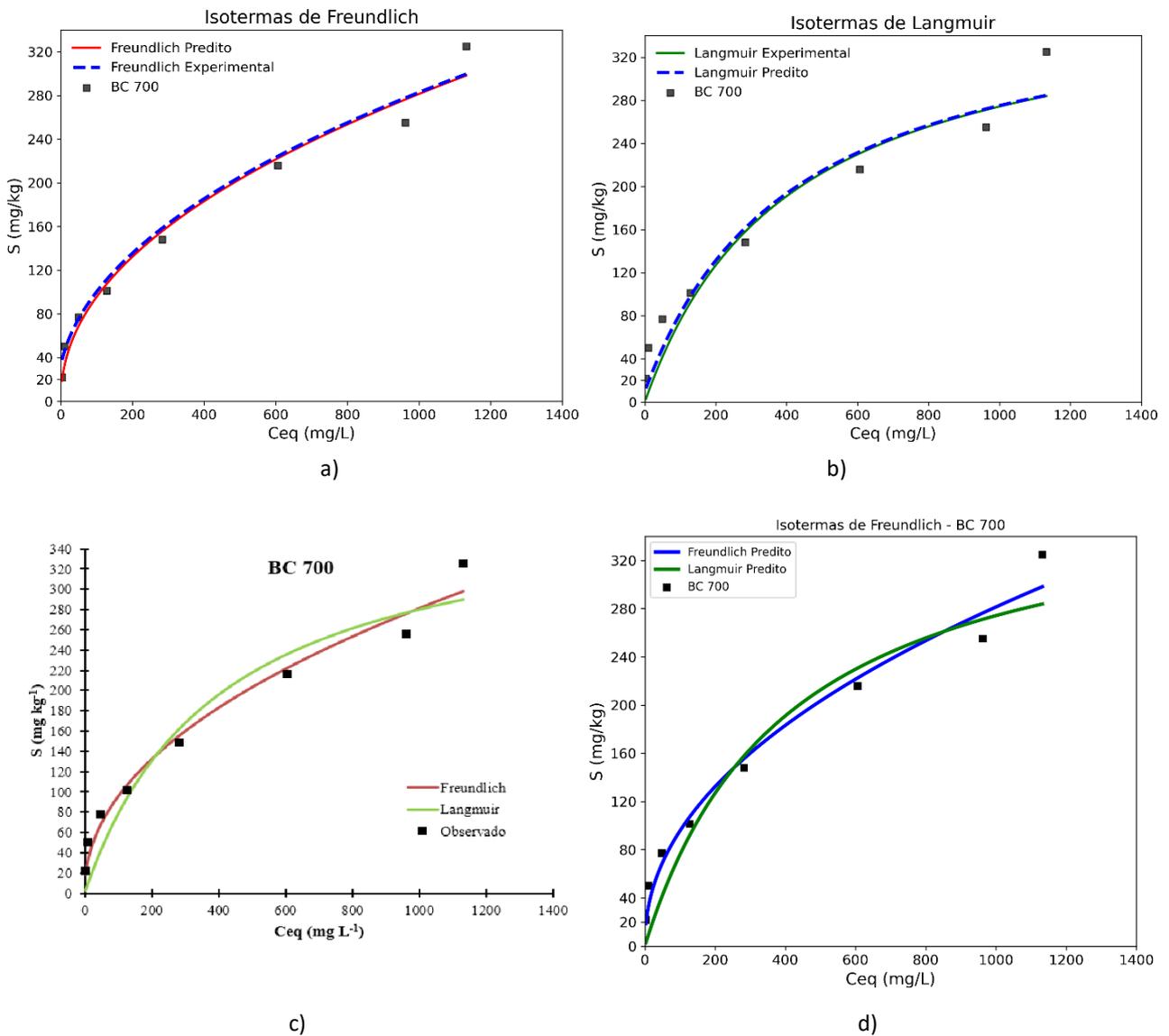
Tabela 2: Valores de AIC e BIC baseados na estimativa da DenseNet.

Modelo	AIC	BIC	AIC	BIC
	Penalização		Ajuste matemático	
Freundlich	-9,59	-2,88	128,02	134,73
Langmuir	-8,27	-1,57	128,68	135,39

Fonte: Autor (2024)

A isoterma de Freundlich estimada apresentou menores valores de AIC e BIC, indicando ser a melhor para descrever o processo adsorptivo do biocarvão, quando estimado com E_{Tr} e K_s , condizente com os gráficos e resultados experimentais.

Figura 13 – Análises comparativas entre as isotermas experimentais e estimadas – a) Freundlich originais e estimadas; b) Langmuir originais e estimadas; c) Freundlich vs Langmuir experimentais; Freundlich vs Langmuir estimadas.



Fonte: Autor (2024)

Como aprimoramento desta pesquisa, novas investigações poderão ser desenvolvidas, como melhorar a configuração dos modelos neurais, combinar modelos em arquiteturas híbridas, gerar variáveis preditoras na mesma área de pesquisa, aumentando a precisão da estimativa, investigar outras técnicas de análises explicativas, melhorando entendimento e interpretação das variáveis envolvidas, avaliar outras isotermas para melhorar efeitos de

energia heterogênea no sistema, integrar a otimização das estimativas, via aprendizado de máquina, estimando diretamente os parâmetros das isotermas.

Direções futuras para ampliação desta pesquisa poderão incluir modelagens de experimentos ainda complexos, como a atividade redox do biocarvão, a influência de agrotóxicos multissítios, cada vez mais frequentes, em solos, quando remediados com biocarvões; a contribuição da modelagem inteligente para redução de custos produtivos do biocarvão em larga escala; ampliação de pesquisas no solo considerando diversas variáveis climáticas; pode ainda, aumentar a precisão dos modelos preditivos e otimizar os processos de produção; além disso, contribui para pesquisas de desenvolvimento de sistemas aperfeiçoados de agricultura de precisão com biocarvão, hidrologia, no contexto de processos ambientais.

6.4 CONCLUSÕES

A modelagem baseada na DenseNet demonstrou alto desempenho preditivo, frente a outros modelos de aprendizado de máquina convencionais, com baixos erros, conseguindo explicar as correlações e influência das variáveis preditoras na estimativa da capacidade adsorptiva, inerentes ao espectro de atuação do biocarvão.

No experimento original, a isoterma de Freundlich seguiu um delineamento não linear, característica de sistemas heterogêneos, onde geralmente ocorre o evento da adsorção, nas regiões de espectro com diferentes energias, conseguindo melhor representar o resultado para o biocarvão produzido a 700°C.

Nos valores estimados pela DenseNet, a curva representativa da isoterma de Freundlich ajustou melhor a distribuição dos dados do biocarvão, com inclusão das duas variáveis (E_{Tr} e K_s), mesmo com aumento da complexidade do sistema em múltiplas regiões de adsorção e maior variedade de energias de ligação, sendo seu resultado ligeiramente superior ao resultado da Freundlich experimental. Isso indica que o modelo aqui proposto pode ainda ser refinado, para melhores aproximações.

A aplicação combinada de redes neurais profundas, estruturadas em uma hierarquia ensemble, confirmou seu alto poder para capturar as relações não lineares complexas entre as variáveis, amparadas por técnicas de análises explicativas eficientes, capazes de interpretar com boa precisão, sua participação na predição, indicando a possibilidade de

novas pesquisas, envolvendo novos atributos, variáveis e métodos produtivos do biocarvão, auxiliando na compreensão de processos de difícil modelagem.

A combinação de redes neurais profundas e modelo ensemble, técnicas robustas, como análises explicativas (ALE, SHAP, IG, Matriz Correlação) e estatísticas (AIC/BIC), permite não apenas prever a adsorção de forma precisa, mas também avaliar a adequação dos modelos isotérmicos, trazendo novas perspectivas para o desenvolvimento de processos ambientais.

REFERÊNCIAS

- ABHAYASIMHA K. C., RAO, C. S., & NAIR, V. Combination of ensemble machine learning models in photocatalytic studies using nano tio₂ - lignin based biochar, **Chemosphere**, volume 352, 141326, issn 0045-6535, 2024.
- ADENIYI, A. G., IWUOZOR, K. O., EMENIKE, E. C., AMOLOYE, M. A., ADELEKE, J. A., OMONAYIN, E. O., BAMIGBOLA, J. O., OJO, H. T., & EZZAT, A. O. Leaf-based biochar: A review of thermochemical conversion techniques and properties. **Journal of Analytical and Applied Pyrolysis**, v. 177, p. 106352, jan. 2024.
- HAMA AZIZ, K. H.; FATAH, N. M.; MUHAMMAD, K. T. Advancements in application of modified biochar as a green and low-cost adsorbent for wastewater remediation from organic dyes. **Royal Society Open Science**, v. 11, n. 5, p. 232033, maio 2024.
- CHEN, Y., ZHAN, B., GUO, B., HONGCEN, L., TIAN, D., DAI, S., YE, P., QIN, H., GAO, P., & YU, Q. Accelerated carbonation curing of biochar-cement mortar: effects of biochar pyrolysis temperatures on carbon sequestration, **Mechanical Properties and Microstructure, Construction and Building Materials**, volume 449, 138446, 2024.
- EBRAHIMI, M., GHOLIPOUR, S., MOSTAFALI, G., & YOUSEFIAN, F. Biochar-amended food waste compost: A review of properties. **Results in Engineering**, v. 24, p. 103118, dez. 2024.
- GABSI, A.E.H., BEN AISSA, C. & MATHLOUTHI, S. A comparative study of basic and ensemble artificial intelligence models for surface roughness prediction during the aa7075 milling process. **Int J Adv Manuf Technol** 126, 1–15 (2023).
- GHANBARI, F., KAMALAN, H. & SARRAF, A. Predicting solid waste generation based on the ensemble artificial intelligence models under uncertainty analysis. **J Mater Cycles Waste Manag** 25, 920–930 (2023).
- GOTORE, O.; MASERE, T. P.; MURONDA, M. T. The immobilization and adsorption mechanisms of agro-waste based biochar: A review on the effectiveness of pyrolytic temperatures on heavy metal removal. **Environmental Chemistry and Ecotoxicology**, v. 6, p. 92–103, 2024.

- GOU, J., SAJID, G. H., SABRI, M. M., EL-MELIGY, K., & OTHMAN, N. A. Optimizing biochar yield and composition prediction with ensemble machine learning models for sustainable production, **Ain Shams Engineering Journal**, volume 16, issue 1, 103209, 2025
- GUPTA, D.; DAS, A.; MITRA, S. Role of modeling and artificial intelligence in process parameter optimization of biochar: A review. **Bioresource Technology**, v. 390, p. 129792, dez. 2023.
- HUANG, J., TAN, X., ALI, I., XIE, Y., & DUAN, Z. Exploring the effect of ulva prolifera components on the biochar carbon sequestration potential, **Journal of Cleaner Production**, volume 484, 144373, 2024
- KANG, W., TIAN, J., REEMT BOGENA, H., LAI, Y., XUE, D., & HE, C. Soil moisture observations and machine learning reveal preferential flow mechanisms in the Qilian Mountains. **Geoderma**, v. 438, p. 116626, out. 2023.
- KORKMAZ, M. SoC estimation of lithium-ion batteries based on machine learning techniques: A filtered approach. **Journal of Energy Storage**, v. 72, p. 108268, nov. 2023.
- LI, X., LIU, W., ZHANG, J., WANG, Z., GUO, Z., ALI, J., WANG, L., YU, Z., ZHANG, X., & SUN, Y. Effective removal of microplastics by filamentous algae and its magnetic biochar: Performance and mechanism. **Chemosphere**, v. 358, p. 142152, jun. 2024.
- LI, Y.; ZHONG, J.; CHANG, L. Characteristics and Drivers of Soil Ecological Stoichiometry in Saline–Alkali Areas of Western Jilin Province, Northeast China. **Land Degradation & Development**, v. 35, n. 17, p. 5411–5424, nov. 2024.
- LV, M. Y.; YU, H. X.; SHANG, X. Y. Sludge-derived biochar: A review on the influence of synthesis conditions on environmental risk reduction and removal mechanism of wastewater pollutants. **Archives of Environmental Protection**, 29 nov. 2023.
- MA, J., ZHANG, S., LIU, X., & WANG, J. Machine learning prediction of biochar yield based on biomass characteristics. **Bioresource Technology**, volume 389, 129820, 2022.
- MITTAL, S., WAZID, M., SINGH, D. P., DAS, A. K., & HOSSAIN, M. S. A deep learning ensemble approach for malware detection in internet of things utilizing explainable artificial intelligence, **Engineering Applications of Artificial Intelligence**, volume 139, part a, 109560, 2025.
- MOZAFFARI, H.; MOOSAVI, A. A.; NEMATOLLAHI, M. A. Predicting saturated and near-saturated hydraulic conductivity using artificial neural network and multiple linear regression in calcareous soils. **PLOS ONE**, v. 19, n. 1, p. e0296933, 10 jan. 2024.
- MURTAZA, G., AHMED, Z., USMAN, M., IQBAL, R., ZULFIQAR, F., TARIQ, A., & DITTA, A. Physicochemical properties and performance of non-woody derived biochars for the sustainable removal of aquatic pollutants: A systematic review. **Chemosphere**, v. 359, p. 142368, jul. 2024.
- OBEY, G.; ADELAIDE, M.; RAMARAJ, R. Biochar derived from non-customized matamba fruit shell as an adsorbent for wastewater treatment. **Journal of Bioresources and Bioproducts**, v. 7, n. 2, p. 109–115, maio 2022.
- PALANSOORIYA, K., LI, J., DISSANAYAKE, P., SUVARNA, M., LI, L., YUAN, X., SARKAR, B., TSANG, D., RINKLEBE, J., WANG, X., & OK, Y. Prediction of soil heavy metal immobilization by biochar using machine learning. **Environmental Science & Technology**, 56, 4187 - 4198, 2022.
- REX, P., MOHAMMED ISMAIL, K., MEENAKSHISUNDARAM, N., BARMAVATU, P., & SAI BHARADWAJ, A. Agricultural Biomass Waste to Biochar: A Review on Biochar Applications

Using Machine Learning Approach and Circular Economy. **ChemEngineering**, v. 7, n. 3, p. 50, 28 de maio 2023.

SHAHEEN, S. M., ANTONIADIS, V., SHAHID, M., YANG, Y., ABDELRAHMAN, H., ZHANG, T., HASSAN, N. E. E., BIBI, I., NIAZI, N. K., YOUNIS, S. A., ALMAZROUI, M., TSANG, Y. F., SARMAH, A. K., KIM, K.-H., & RINKLEBE, J. Sustainable applications of rice feedstock in agro-environmental and construction sectors: A global perspective. **Renewable and Sustainable Energy Reviews**, v. 153, p. 111791, jan. 2022.

SINGH, K., SINGH, B., SIHAG, P., KUMAR, V., & SHARMA, K. V. Development and application of modeling techniques to estimate the unsaturated hydraulic conductivity. **Modeling Earth Systems and Environment**, v. 9, n. 4, p. 4557–4571, nov. 2023.

SONG, Q., KONG, F., LIU, B.-F., SONG, X., & REN, H.-Y. Biochar-based composites for removing chlorinated organic pollutants: Applications, mechanisms, and perspectives. **Environmental Science and Ecotechnology**, v. 21, p. 100420, set. 2024.

SU, G., JIANG, & P. Machine learning models for predicting biochar properties from lignocellulosic biomass torrefaction, **Bioresource Technology**, volume 399, 130519, 2024.

TANG, J. Y., CHUNG, B. Y. H., ANG, J. C., CHONG, J. W., TAN, R. R., AVISO, K. B., CHEMMANGATTUVALAPPIL, N. G., & THANGALAZHY-GOPAKUMAR, S. Prediction model for biochar energy potential based on biomass properties and pyrolysis conditions derived from rough set machine learning. **Environmental Technology**, v. 45, n. 15, p. 2908–2922, 2 jul. 2024.

WEN, R.; LI, S. Spatial Decision Support Systems with Automated Machine Learning: A Review. **ISPRS International Journal of Geo-Information**, v. 12, n. 1, p. 12, 30 dez. 2022.

XIANG, Y., LIU, Y., NIAZI, N. K., BOLAN, N., ZHAO, L., ZHANG, S., XUE, J., YAO, B., & LI, Y. Biochar addition increased soil bacterial diversity and richness: Large-scale evidence of field experiments. **Science of The Total Environment**, v. 893, p. 164961, out. 2023.

XIE, G., ZHU, C., LI, C., FAN, Z., & WANG, B. Predicting the adsorption of ammonia nitrogen by biochar in water bodies using machine learning strategies: model optimization and analysis of key characteristic variables, **Environmental Research**, volume 267, 2025.

XIE, J., LATIF, J., YANG, K., WANG, Z., ZHU, L., YANG, H., QIN, J., NI, Z., JIA, H., XIN, W., & LI, X. A state-of-art review on the redox activity of persistent free radicals in biochar. **Water Research**, v. 255, p. 121516, maio 2024.

YAO, C., WANG, B., ZHANG, J., FAHEEM, M., FENG, Q., HASSAN, M., ZHANG, X., LEE, X., & WANG, S. Formation mechanisms and degradation methods of polycyclic aromatic hydrocarbons in biochar: A review. **Journal of Environmental Management**, v. 357, p. 120610, abr. 2024.

YAQUB, M., & LEE, W. Artificial intelligence models for predicting calcium and magnesium removal by polyfunctional ketone using ensemble machine learners, **Chemosphere**, volume 345, 140422, 2023.

YIN, R., LI, X., NING, Y., HU, Q., MAO, Y., ZHANG, X., & ZHANG, X. Machine learning unveils the role of biochar application in enhancing tea yield by mitigating soil acidification in tea plantations, **Science of the Total Environment**, volume 965, 2025.

ZHANG, T., TANG, Y., LI, H., HU, W., CHENG, J., & LEE, X. A bibliometric review of biochar for soil carbon sequestration and mitigation from 2001 to 2020. **Ecotoxicology and Environmental Safety**, v. 264, p. 115438, out. 2023.

ZHENG, H., CHEN, Q., TANG, B., ZHAO, M., LU, G., WANG, J., & SONG, G. Effectiveness evaluation method of marine environmental weapons and equipment based on ensemble

leaning. in: nakamatsu, k., kountchev, r., patnaik, s., abe, j.m., tyugashev, a. (eds) advanced intelligent technologies for industry. smart innovation, **Systems and Technologies**, vol 285. **springer, singapore**, 2025.

ZHU, H., AN, Q., SYAFIKA MOHD NASIR, A., BABIN, A., LUCERO SAUCEDO, S., VALLENAS, A., LI, L., BALDWIN, S. A., LAU, A., & BI, X. Emerging applications of biochar: a review on techno-environmental-economic aspects. **Bioresource Technology**, 2023.

7 CONCLUSÕES GERAIS

A pesquisa acadêmica subsidiou os principais enfoques ligados ao biocarvão, delimitando a área de pesquisa. A compreensão de seus mecanismos de atuação, a modelagem de seus atributos-chave e parâmetros do meio são representados por variáveis complexas, de alta variabilidade espacial, com interação não linear, exigindo esforço estatístico e computacional para sua predição.

Definir quais modelos de inteligência artificial melhor retratam o delineamento das variáveis passa necessariamente por análise prévia dos dados de entrada ao modelo, definindo suas características válidas.

A predominância de pesquisas com biocarvão envolvendo modelagem com aprendizado de máquina tem focado na otimização de resultados experimentais, na pesquisa com sua biomassa, ou mesmo de seus parâmetros produtivos. A disponibilidade de poderosos algoritmos de redes neurais profundas, como a DenseNet, expandindo pequenos conjuntos de dados com algoritmos como WGAN, ou mesmo pela aquisição indireta, como de imageamento, abre novas perspectivas de pesquisas, apoiados por ferramentas e técnicas de análises explicativas, como SHAP, LIME, matriz de correlação, associadas, auxiliando na compreensão da atuação e influência das variáveis envolvidas no processo de modelagem.

Neste trabalho, a aplicação de uma hierarquia de três redes neurais profundas (ensemble), com duas estimando dados para integração a outros disponíveis, permitiu que uma terceira rede avaliasse a capacidade adsorptiva de um biocarvão para imobilizar moléculas de Imazapic, em experimento de bancada. A modelagem apresentou resultado semelhante, otimizando a capacidade adsorptiva do experimento em aproximadamente 3%, comprovando a viabilidade do experimento, abrindo caminho para ampliação desta pesquisa, além de outras pesquisas nesta direção.

REFERÊNCIAS

- ABBA, S. I.; BENAIFI, M.; ALJUNDI, I. H. Intelligent process optimisation based on cutting-edge emotional learning for performance evaluation of NF/RO of seawater desalination plant. **Desalination**, v. 550, p. 116376, mar. 2023.
- ABHAYASIMHA K. C., RAO, C. S., & NAIR, V. Combination of ensemble machine learning models in photocatalytic studies using nano tio₂ - lignin based biochar, **Chemosphere**, volume 352, 141326, issn 0045-6535, 2024.
- ADENIYI, A. G., IWUOZOR, K. O., EMENIKE, E. C., AMOLOYE, M. A., ADELEKE, J. A., OMONAYIN, E. O., BAMIGBOLA, J. O., OJO, H. T., & EZZAT, A. O. Leaf-based biochar: A review of thermochemical conversion techniques and properties. **Journal of Analytical and Applied Pyrolysis**, v. 177, p. 106352, jan. 2024.
- AHMADISHARAF, A. et al. Representative Sample Size for Estimating Saturated Hydraulic Conductivity via Machine Learning: A Proof-Of-Concept Study. **Water Resources Research**, v. 60, n. 8, p. e2023WR036783, ago. 2024.
- AL MAMUN, MD. A., SARKER, M. R., SARKAR, M. A. R., ROY, S. K., NIHAD, S. A. I., MCKENZIE, A. M., HOSSAIN, MD. I., & KABIR, MD. S. Identification of influential weather parameters and seasonal drought prediction in Bangladesh using machine learning algorithm. **Scientific Reports**, v. 14, n. 1, p. 566, 4 jan. 2024.
- ALBALASMEH, A. et al. Artificial neural network optimization to predict saturated hydraulic conductivity in arid and semi-arid regions. **CATENA**, v. 217, p. 106459, out. 2022.
- ALSENJAR, O., CETIN, M., AKSU, H. ET AL. Actual evapotranspiration estimation using metric model and landsat satellite images over an irrigated field in the eastern mediterranean region of turkey. **Med. Geosc. rev.** 5, 35–49 (2023). <https://doi.org/10.1007/s42990-023-00099-y>
- BAFTI, A. G., AHMADI, A., ABBASI, A., KAMANGIR, H., JAMALI, S., & HASHEMI, H. Automated actual evapotranspiration estimation: Hybrid model of a novel attention based U-Net and metaheuristic optimization algorithms. **Atmospheric Research**, v. 297, p. 107107, jan. 2024.
- BAŞAKIN, E. E.; EKMEKCIOĞLU, Ö.; ÖZGER, M. Providing a comprehensive understanding of missing data imputation processes in evapotranspiration-related research: a systematic literature review. **Hydrological Sciences Journal**, v. 68, n. 14, p. 2089–2104, 26 out. 2023.
- BECK, H. E., ZIMMERMANN, N. E., MCVICAR, T. R., VERGOPOLAN, N., BERG, A., & WOOD, E. F. (2018). Present and future köppen-geiger climate classification maps at 1-km resolution. **Scientific Data**, 5(1), 180214. <https://doi.org/10.1038/sdata.2018.214>, 2018.
- BEYER, M., AHMAD, R., YANG, B., & RODRÍGUEZ-BOCCA, P. Deep spatial-temporal graph modeling for efficient NDVI forecasting. **Smart Agricultural Technology**, v. 4, p. 100172, ago. 2023.
- BURCHARD-LEVINE, V., BORRA-SERRANO, I., PEÑA, J. M., KUSTAS, W. P., GUERRA, J. G., DORADO, J., MESÍAS-RUIZ, G., HERREZUELO, M., MARY, B., MCKEE, L. M., DE CASTRO, A. I., SANCHEZ-ÉLEZ, S., & NIETO, H. Evaluating the precise grapevine water stress detection using unmanned aerial vehicles and evapotranspiration-based metrics. **Irrigation Science**, 29 abr. 2024.
- BUTLER, R. A., PAPEŞ, M., VOGT, J. T., PAULSEN, D. J., CROWE, C., & TROUT FRYXELL, R. T. Human risk to tick encounters in the southeastern United States estimated

with spatial distribution modeling. **PLOS Neglected Tropical Diseases**, v. 18, n. 2, p. e0011919, 14 fev. 2024.

CHAKHAR, A., ZITOUNA-CHEBBI, R., HERNÁNDEZ-LÓPEZ, D., BALLESTEROS, R., MAHJOUB, I., & MORENO, M. A. Assessing the Accuracy of Multiple Classification Algorithms Combining Sentinel-1 and Sentinel-2 for the Citrus Crop Classification and spatialization of the Actual Evapotranspiration Obtained from Flux Tower Eddy Covariance: Case Study of Cap Bon, Tunisia. **Proceedings of IAHS**, v. 385, p. 443–448, 19 abr. 2024.

CHANG, F., YANG, W., WANG, S., YIN, L., & DENG, X. Long-term optimization of agronomic practices increases water storage capacity and available water in soil. **Agronomy**, 14(10), 2286, 2024.

CHEN, B.; LI, J.; JIN, Y. Deep Learning for Feature-Level Data Fusion: Higher Resolution Reconstruction of Historical Landsat Archive. **Remote Sensing**, v. 13, n. 2, p. 167, 6 jan. 2021.

CHEN, Y., ZHAN, B., GUO, B., HONGCEN, L., TIAN, D., DAI, S., YE, P., QIN, H., GAO, P., & YU, Q. Accelerated carbonation curing of biochar-cement mortar: effects of biochar pyrolysis temperatures on carbon sequestration, **Mechanical Properties and Microstructure, Construction and Building Materials**, volume 449, 138446, 2024.

CHOUDHARY, C. S. et al. Mechanisms of allelopathic interactions for sustainable weed management. **Rhizosphere**, v. 25, p. 100667, mar. 2023.

DA SILVA, E. H. F. M., HOOGENBOOM, G., BOOTE, K. J., GONÇALVES, A. O., & MARIN, F. R. Predicting soybean evapotranspiration and crop water productivity for a tropical environment using the CSM-CROPGRO-Soybean model. **Agricultural and Forest Meteorology**, v. 323, p. 109075, ago. 2022.

DANDA, T. J., KUSANGAYA, S., MUSHORE, T. D., & MUJERE, N. Assessing the role of evapotranspiration in reducing surface temperatures in Harare using the SEBAL algorithm. **Urban Climate**, v. 49, p. 101442, maio 2023.

DARAMOLA, O. S., MacDONALD, E., KANISSERY, R. G., TILLMAN, B. L., SINGH, H., AJANI, O. A. A., DEVKOTA, P. Effect of planting pattern and herbicide programs on sicklepod (*Senna obtusifolia* L.) control in peanut. **Weed Technology**, v. 38, p. e53, 2024.

DAYAL, D., ET AL. Streamflow estimation using satellite-retrieved water fluxes and machine learning technique over monsoon-dominated catchments of India. **Hydrol. Sci. J.** 66 (4), 656–671. <https://doi.org/10.1080/02626667.2021.1889557>, 2021.

DE RIJK, V. et al. **Predicting Saturated Hydraulic Conductivity from Particle Size Distributions Using Machine Learning**. SSRN, 2024. Disponível em: <<https://www.ssrn.com/abstract=4862913>>. Acesso em: 16 nov. 2024

DE SOUZA, C. A. A., JARDIM, A. M. D. R. F., DE SOUZA, L. S. B., JÚNIOR, G. D. N. A., ALVES, C. P., DE MORAIS, J. E. F., LOPES, D. D. C., STEIDLE NETO, A. J., SALVADOR, K. R. D. S., & DA SILVA, T. G. F. Intercomparison of micrometeorological variables, surface energy fluxes, and evapotranspiration in different landscapes of the Brazilian semi-arid region. **Agricultural and Forest Meteorology**, v. 341, p. 109679, out. 2023.

DE-VILLE, S., EDMONDSON, J., GREEN, D., STIRLING, R., DAWSON, R., & STOVIN, V. Effect of vegetation treatment and water stress on evapotranspiration in bioretention systems. **Water Research**, v. 252, p. 121182, mar. 2024.

DIYKH, M., ALI, M., JAMEI, M., ABDULLA, S., UDDIN, M. P., FAROOQUE, A. A., LABBAN, A. H., & ALABDALLY, H. Empirical curvelet transform based deep DenseNet model to

predict NDVI using RGB drone imagery data. **Computers and Electronics in Agriculture**, v. 221, p. 108964, jun. 2024.

DOUNA, V., BARRAZA, V., GRINGS, F., HUETE, A., RESTREPO-COUBE, N., & BERINGER, J. Towards a remote sensing data based evapotranspiration estimation in Northern Australia using a simple random forest approach. **Journal of Arid Environments**, v. 191, p. 104513, ago. 2021.

EBRAHIMI, M., GHOLIPOUR, S., MOSTAFALI, G., & YOUSEFIAN, F. Biochar-amended food waste compost: A review of properties. **Results in Engineering**, v. 24, p. 103118, dez. 2024.

ELZAIN, H. E., ABDALLA, O. A., ABDALLAH, M., AL-MAKTOUMI, A., ELTAYEB, M., & ABBA, S. I. Innovative approach for predicting daily reference evapotranspiration using improved shallow and deep learning models in a coastal region: A comparative study. **Journal of Environmental Management**, v. 354, p. 120246, mar. 2024.

FAE, G. S., KEMANIAN, A. R., ROTH, G. W., WHITE, C., & WATSON, J. E. Soybean yield in relation to environmental and soil properties. **European Journal of Agronomy**, 118, 126070, 2020.

FANG, Y., XU, L., CHEN, Y., ZHOU, W., WONG, A., & CLAUSI, D. A. A Bayesian Deep Image Prior Downscaling Approach for High-Resolution Soil Moisture Estimation. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 15, p. 4571–4582, 2022.

FAOUZI, E. et al. Spatial mapping of hydrologic soil groups using machine learning in the Mediterranean region. **CATENA**, v. 232, p. 107364, nov. 2023.

GABSI, A.E.H., BEN AISSA, C. & MATHLOUTHI, S. A comparative study of basic and ensemble artificial intelligence models for surface roughness prediction during the aa7075 milling process. **Int J Adv Manuf Technol** 126, 1–15 (2023).

GEE. GOOGLE EARTH ENGINE, PLATFORM. **Produto GEE** Disponível em: <https://earthengine.google.com/>. Acesso em: 22 set. 2024.

GHANBARI, F., KAMALAN, H. & SARRAF, A. Predicting solid waste generation based on the ensemble artificial intelligence models under uncertainty analysis. **J Mater Cycles Waste Manag** 25, 920–930 (2023).

GONG, H., CHENG, Q., JIN, H., & REN, Y. Effects of temporal, spatial, and elevational variation in bioclimatic indices on the NDVI of different vegetation types in Southwest China. **Ecological Indicators**, v. 154, p. 110499, out. 2023.

GONZALES-INCA, C. et al. Geospatial Artificial Intelligence (GeoAI) in the Integrated Hydrological and Fluvial Systems Modeling: Review of Current Applications and Trends. **Water**, v. 14, n. 14, p. 2211, 13 jul. 2022.

GONZÁLEZ PEREA, R., FERNÁNDEZ GARCÍA, I., CAMACHO POYATO, E., & RODRÍGUEZ DÍAZ, J. A. New memory-based hybrid model for middle-term water demand forecasting in irrigated areas. **Agricultural Water Management**, v. 284, p. 108367, jun. 2023.

GOTORE, O.; MASERE, T. P.; MURONDA, M. T. The immobilization and adsorption mechanisms of agro-waste based biochar: A review on the effectiveness of pyrolytic temperatures on heavy metal removal. **Environmental Chemistry and Ecotoxicology**, v. 6, p. 92–103, 2024.

- GOU, J., SAJID, G. H., SABRI, M. M., EL-MELIGY, K., & OTHMAN, N. A. Optimizing biochar yield and composition prediction with ensemble machine learning models for sustainable production, **Ain Shams Engineering Journal**, volume 16, issue 1, 103209, 2025
- GUO, H., HU, S., WANG, X., CHANG, M.-C., & LYU, S. Robust Attentive Deep Neural Network for Detecting GAN-Generated Faces. **IEEE Access**, v. 10, p. 32574–32583, 2022.
- GUO, Y., ZHANG, L., HE, Y., CAO, S., LI, H., RAN, L., DING, Y., FILONCHYK, M. LSTM time series ndvi prediction method incorporating climate elements: a case study of yellow river basin. China. **Journal of Hydrology** 629, 130518, 2024. <https://doi.org/10.1016/j.jhydrol.2023.130518>, 2023.
- GUPTA, D.; DAS, A.; MITRA, S. Role of modeling and artificial intelligence in process parameter optimization of biochar: A review. **Bioresource Technology**, v. 390, p. 129792, dez. 2023.
- HAMA AZIZ, K. H.; FATAH, N. M.; MUHAMMAD, K. T. Advancements in application of modified biochar as a green and low-cost adsorbent for wastewater remediation from organic dyes. **Royal Society Open Science**, v. 11, n. 5, p. 232033, maio 2024.
- HAO, P., DI, L., & GUO, L. Estimation of crop evapotranspiration from MODIS data by combining random forest and trapezoidal models. **Agricultural Water Management**, v. 259, p. 107249, jan. 2022.
- HE, L., DUAN, X., DING, S., CAI, C., & DENG, Y. Construction and evaluation of pedotransfer functions for saturated hydraulic conductivity in the granite red soil regions of southern China. **Journal of Hydrology: Regional Studies**, v. 50, p. 101539, dez. 2023.
- HUANG, J., TAN, X., ALI, I., XIE, Y., & DUAN, Z. Exploring the effect of ulva prolifera components on the biochar carbon sequestration potential, **Journal of Cleaner Production**, volume 484, 144373, 2024
- JALAL, F. E., IQBAL, M., ALI KHAN, M., SALAMI, B. A., ULLAH, S., KHAN, H., & NABIL, M. Indirect Estimation of Swelling Pressure of Expansive Soil: GEP versus MEP Modelling. **Advances in Materials Science and Engineering**, v. 2023, p. 1–25, 23 jan. 2023.
- JEBRAEELI, V., JIANG, B., KRIM, H., & CANSEVER, D. **Generative Expansion of Small Datasets: An Expansive Graph Approach**. arXiv, 1 out. 2024. Disponível em: <<http://arxiv.org/abs/2406.17238>>. Acesso em: 10 dez. 2024.
- JOZDANI, S., CHEN, D., POULIOT, D., & ALAN JOHNSON, B. A review and meta-analysis of Generative Adversarial Networks and their applications in remote sensing. **International Journal of Applied Earth Observation and Geoinformation**, v. 108, p. 102734, abr. 2022.
- JUNKES, G. V., AVILA, L., KEMMERICH, M., GERKE, V. R., FIPKE, M. V., CAMARGO, E. R. Imidazolinone herbicide dissipation in rice fields as affected by intermittent and continuous irrigation. **Advances in Weed Science**, v. 40, p. e020220094, 20 maio 2022.
- KANG, W., TIAN, J., REEMT BOGENA, H., LAI, Y., XUE, D., & HE, C. Soil moisture observations and machine learning reveal preferential flow mechanisms in the Qilian Mountains. **Geoderma**, v. 438, p. 116626, out. 2023.
- KARUNATHILAKE, E. M. B. M., LE, A. T., HEO, S., CHUNG, Y. S., & MANSOOR, S. The Path to Smart Farming: Innovations and Opportunities in Precision Agriculture. **Agriculture**, v. 13, n. 8, p. 1593, 11 ago. 2023.
- KEÏTA, A., ZOROM, M., FAYE, M. D., DAMBA, D. D., KONATE, Y., HAYDE, L. G., & LIDON, B. Achieving Real-World Saturated Hydraulic Conductivity: Practical and Theoretical Findings from Using an Exponential One-Phase Decay Model. **Hydrology**, v. 10, n. 12, p. 235, 9 dez. 2023.

- KHARATE, N., ANERAIO, P., KULKARNI, A., & ABDULLAH, M. Explainable AI Techniques for Comprehensive Analysis of the Relationship between Process Parameters and Material Properties in FDM-Based 3D-Printed Biocomposites. **Journal of Manufacturing and Materials Processing**, v. 8, n. 4, p. 171, 6 ago. 2024.
- KHARATE, N., ANERAIO, P., KULKARNI, A., & ABDULLAH, M. Explainable AI Techniques for Comprehensive Analysis of the Relationship between Process Parameters and Material Properties in FDM-Based 3D-Printed Biocomposites. **Journal of Manufacturing and Materials Processing**, v. 8, n. 4, p. 171, 6 ago. 2024.
- KIM, S., LEE, E., HWANG, H.-T., PYO, J., YUN, D., BAEK, S.-S., & CHO, K. H. Spatiotemporal estimation of groundwater and surface water conditions by integrating deep learning and physics-based watershed models. **Water Research X**, v. 23, p. 100228, maio 2024.
- KIRAN, A.; KUMAR, S. S. A Methodology and an Empirical Analysis to Determine the Most Suitable Synthetic Data Generator. **IEEE Access**, v. 12, p. 12209–12228, 2024.
- KORKMAZ, M. SoC estimation of lithium-ion batteries based on machine learning techniques: A filtered approach. **Journal of Energy Storage**, v. 72, p. 108268, nov. 2023.
- KUMAR, V., KEDAM, N., SHARMA, K. V., MEHTA, D. J., & CALOIERO, T. Advanced Machine Learning Techniques to Improve Hydrological Prediction: A Comparative Analysis of Streamflow Prediction Models. **Water**, v. 15, n. 14, p. 2572, 13 jul. 2023.
- LI, D., JIANG, Z., TIAN, K., JI, R. Prediction of hydraulic conductivity of sodium bentonite gcls by machine learning approaches. **Environmental Geotechnics**, e-issn 2051-803x, (2023), <https://doi.org/10.1680/jenge.22.00181>.
- LI, X., LIU, W., ZHANG, J., WANG, Z., GUO, Z., ALI, J., WANG, L., YU, Z., ZHANG, X., & SUN, Y. Effective removal of microplastics by filamentous algae and its magnetic biochar: Performance and mechanism. **Chemosphere**, v. 358, p. 142152, jun. 2024.
- LI, Y.; ZHONG, J.; CHANG, L. Characteristics and Drivers of Soil Ecological Stoichiometry in Saline–Alkali Areas of Western Jilin Province, Northeast China. **Land Degradation & Development**, v. 35, n. 17, p. 5411–5424, nov. 2024.
- LIU, H., LI, B., WU, H., LIANG, H., HUANG, Y., LI, Y., GHANEM, B., & ZHENG, Y. Combating Mode Collapse via Offline Manifold Entropy Estimation. **Proceedings of the AAAI Conference on Artificial Intelligence**, v. 37, n. 7, p. 8834–8842, 26 jun. 2023.
- LIU, X. "Deep learning in marketing: a review and research agenda", SUDHIR, K. AND TOUBIA, O. (ed.) artificial intelligence in marketing (review of marketing research, vol. 20), **Emerald Publishing Limited**, leeds, pp. 239-271, (2023).
- LÜ, X., NURMEMET I., XIAO, S., ZHAO, J., YU, X., AILI, Y., LI, S. Spatial-temporal simulation and prediction of root zone soil moisture based on hydrus-1d and cnn-lstm-attention in the yutian oasis, southern xinjiang, china, **Pedosphere**, 2024, issn 1002-0160, (2024). <https://doi.org/10.1016/j.pedsph.2024.10.008>.
- LV, M. Y.; YU, H. X.; SHANG, X. Y. Sludge-derived biochar: A review on the influence of synthesis conditions on environmental risk reduction and removal mechanism of wastewater pollutants. **Archives of Environmental Protection**, 29 nov. 2023.
- MA, J., ZHANG, S., LIU, X., & WANG, J. Machine learning prediction of biochar yield based on biomass characteristics. **Bioresource Technology**, vol. 389, 2022.
- MA, Y., SUN, S., LI, C., ZHAO, J., LI, Z., & JIA, C. Estimation of regional actual evapotranspiration based on the improved Sebal model. **Journal of Hydrology**, 619, 129283, (2023).

MAHSA H. KASHANI, MOHAMMAD ALI GHORBANI, MAHMOOD SHAHABI, SUJAY RAGHAVENDRA NAGANNA, LAMINE DIOP. Multiple ai model integration strategy—application to saturated hydraulic conductivity prediction from easily available soil properties. **Soil and Tillage Research**, vol. 196, 2020.

MEHMOOD R, BASHIR R, & GIRI KJ Deep generative models: a review. department of computer science, islamic university of science & technology, kashmir, jammu and kashmir, india. **Indian Journal of Science and Technology**, 16(7), 460–467, 2023.

MITTAL, S., WAZID, M., SINGH, D. P., DAS, A. K., & HOSSAIN, M. S. A deep learning ensemble approach for malware detection in internet of things utilizing explainable artificial intelligence, **Engineering Applications of Artificial Intelligence**, volume 139, part a, 109560, 2025.

MOD16A2.Earth Data, USGS, /Terra Net Evapotranspiration 8-day I4 **Produto Global** 500 m s/n grid. Disponível em: <https://lpdaac.usgs.gov/products/mod16a2v006/>. Acesso em: 19 set. 2024.

MOKHTARI, A., SADEGHI, M., AFRASIABIAN, Y., & YU, K. OPTRAM-ET: A novel approach to remote sensing of actual evapotranspiration applied to Sentinel-2 and Landsat-8 observations. **Remote Sensing of Environment**, v. 286, p. 113443, mar. 2023.

MOZAFFARI, H.; MOOSAVI, A. A.; NEMATOLLAHI, M. A. Predicting saturated and near-saturated hydraulic conductivity using artificial neural network and multiple linear regression in calcareous soils. **PLOS ONE**, v. 19, n. 1, p. e0296933, 10 jan. 2024.

MUÑOZ-CARPENA, R., CARMONA-CABRERO, A., YU, Z., FOX, G., & BATELAAN, O. Convergence of mechanistic modeling and artificial intelligence in hydrologic science and engineering. **PLOS Water**, v. 2, n. 8, p. e0000059, 7 ago. 2023.

MURTAZA, G., AHMED, Z., USMAN, M., IQBAL, R., ZULFIQAR, F., TARIQ, A., & DITTA, A. Physicochemical properties and performance of non-woody derived biochars for the sustainable removal of aquatic pollutants: A systematic review. **Chemosphere**, v. 359, p. 142368, jul. 2024.

OBEY, G.; ADELAIDE, M.; RAMARAJ, R. Biochar derived from non-customized matamba fruit shell as an adsorbent for wastewater treatment. **Journal of Bioresources and Bioproducts**, v. 7, n. 2, p. 109–115, maio 2022.

PALANSOORIYA, K., LI, J., DISSANAYAKE, P., SUVARNA, M., LI, L., YUAN, X., SARKAR, B., TSANG, D., RINKLEBE, J., WANG, X., & OK, Y. Prediction of soil heavy metal immobilization by biochar using machine learning. **Environmental Science & Technology**, 56, 4187 - 4198, 2022.

PULLA, S. T.; YASARER, H.; YARBROUGH, L. D. Synthetic Time Series Data in Groundwater Analytics: Challenges, Insights, and Applications. **Water**, v. 16, n. 7, p. 949, 25 mar. 2024.

QGIS. **Produto QGIS Software**. Disponível em: <https://www.qgis.org/download/>. Acesso em: 25 set. 2024.

RAHIMPOUR, M., RAHIMZADEGAN, M. Assessment of surface energy balance algorithm for land and operational simplified surface energy balance algorithm over freshwater and saline water bodies in urmia lake basin. *Theor. Appl. Climatol.* 143 (3–4), 1457–1472, 2021.

REX, P., MOHAMMED ISMAIL, K., MEENAKSHISUNDARAM, N., BARMAVATU, P., & SAI BHARADWAJ, A. Agricultural Biomass Waste to Biochar: A Review on Biochar Applications

Using Machine Learning Approach and Circular Economy. **ChemEngineering**, v. 7, n. 3, p. 50, 28 de maio 2023.

REZAEI, M., MOUSAVI, S. R., RAHMANI, A., ZERAATPISHEH, M., RAHMATI, M., PAKPARVAR, M., JAHANDIDEH MAHJENABADI, V. A., SEUNTJENS, P., & CORNELIS, W. Incorporating machine learning models and remote sensing to assess the spatial distribution of saturated hydraulic conductivity in a light-textured soil. **Computers and Electronics in Agriculture**, v. 209, p. 107821, jun. 2023.

REZAEI, M., MOUSAVI, S. R., RAHMANI, A., ZERAATPISHEH, M., RAHMATI, M., PAKPARVAR, M., JAHANDIDEH MAHJENABADI, V. A., SEUNTJENS, P., & CORNELIS, W. Incorporating machine learning models and remote sensing to assess the spatial distribution of saturated hydraulic conductivity in a light-textured soil. **Computers and Electronics in Agriculture**, v. 209, p. 107821, jun. 2023.

ROBERT, C. P.; ROUSSEAU, J. A special issue on Bayesian inference: challenges, perspectives and prospects. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 381, n. 2247, p. 20220155, 15 de maio 2023.

SARTORI, H. L., COSTA, A. G. F., SOFIATTI, V., ALVES, P. L. C. A. Application of herbicides and sugar cane straw on controlling of *Mucuna aterrima* L. in peanut crop. **Australian Journal of Crop Science**, n. 17(11):2023, p. 835–840, 1 nov. 2023.

SALIH, A., RAISI-ESTABRAGH, Z., GALAZZO, I. B., RADEVA, P., PETERSEN, S. E., MENEGAZ, G., & LEKADIR, K. A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME. **Advanced Intelligent Systems**, p. 2400304, 27 jun. 2024.

SENAY, G. B., PARRISH, G. E. L., SCHAUER, M., FRIEDRICH, M., KHAND, K., BOIKO, O., KAGONE, S., DITTMEIER, R., ARAB, S., & JI, L. Improving the Operational Simplified Surface Energy Balance Evapotranspiration Model Using the Forcing and Normalizing Operation. **Remote Sensing**, v. 15, n. 1, p. 260, 1 jan. 2023.

SHAHEEN, S. M., ANTONIADIS, V., SHAHID, M., YANG, Y., ABDELRAHMAN, H., ZHANG, T., HASSAN, N. E. E., BIBI, I., NIAZI, N. K., YOUNIS, S. A., ALMAZROUI, M., TSANG, Y. F., SARMAH, A. K., KIM, K.-H., & RINKLEBE, J. Sustainable applications of rice feedstock in agro-environmental and construction sectors: A global perspective. **Renewable and Sustainable Energy Reviews**, v. 153, p. 111791, jan. 2022.

SHAN, W.; LIU, J.; ZHOU, J. Integrated method for intelligent structural design of steel frames based on optimization and machine learning algorithm. **Engineering Structures**, v. 284, p. 115980, jun. 2023.

SHI, S., TSE, R., LUO, W., D'ADDONA, S., & PAU, G. Machine learning-driven credit risk: a systemic review. **Neural Computing and Applications**, v. 34, n. 17, p. 14327–14339, set. 2022.

SHREIM, H.; GIZZINI, A. K.; GHANDOUR, A. J. **Trainable Noise Model as an Explainable Artificial Intelligence Evaluation Method: Application on Sobol for Remote Sensing Image Segmentation**. ECRS 2023. **Anais**. Em: ECRS 2023. MDPI, 6 nov. 2023. Disponível em: <<https://www.mdpi.com/2673-4931/29/1/49>>. Acesso em: 12 dez. 2024.

SINGH, K., SINGH, B., SIHAG, P., KUMAR, V., & SHARMA, K. V. Development and application of modeling techniques to estimate the unsaturated hydraulic conductivity. **Modeling Earth Systems and Environment**, v. 9, n. 4, p. 4557–4571, nov. 2023.

SINGH, V. K. et al. Novel Genetic Algorithm (GA) based hybrid machine learning-pedotransfer Function (ML-PTF) for prediction of spatial pattern of saturated hydraulic

conductivity. **Engineering Applications of Computational Fluid Mechanics**, v. 16, n. 1, p. 1082–1099, 31 dez. 2022.

SLATER, L. J., ARNAL, L., BOUCHER, M.-A., CHANG, A. Y.-Y., MOULDS, S., MURPHY, C., NEARING, G., SHALEV, G., SHEN, C., SPEIGHT, L., VILLARINI, G., WILBY, R. L., WOOD, A., AND ZAPPA, M.: Hybrid forecasting: blending climate predictions with ai models, hydrol. **Earth Syst. Sci.**, 27, 1865–1889, 2023.

SONG, Q., KONG, F., LIU, B.-F., SONG, X., & REN, H.-Y. Biochar-based composites for removing chlorinated organic pollutants: Applications, mechanisms, and perspectives. **Environmental Science and Ecotechnology**, v. 21, p. 100420, set. 2024.

SU, G., JIANG, & P. Machine learning models for predicting biochar properties from lignocellulosic biomass torrefaction, **Bioresource Technology**, volume 399, 130519, 2024.

TAN, W., LIU, P., LI, X., XU, S., CHEN, Y., & YANG, J. Segmentation of lung airways based on deep learning methods. **IET Image Processing**, v. 16, n. 5, p. 1444–1456, abr. 2022.

TANG, J. Y., CHUNG, B. Y. H., ANG, J. C., CHONG, J. W., TAN, R. R., AVISO, K. B., CHEMMANGATTUVALAPPIL, N. G., & THANGALAZHY-GOPAKUMAR, S. Prediction model for biochar energy potential based on biomass properties and pyrolysis conditions derived from rough set machine learning. **Environmental Technology**, v. 45, n. 15, p. 2908–2922, 2 jul. 2024.

TARWIDI, D., PUDJAPRASETYA, S. R., ADYTIA, D., & APRI, M. An optimized XGBoost-based machine learning method for predicting wave run-up on a sloping beach. **MethodsX**, v. 10, p. 102119, 2023.

TREJO-ALONSO, J., FUENTES, S., MORALES-DURAN, N., & CHAVEZ, C. Saturated Hydraulic Conductivity Estimation Using Artificial Neural Networks. **Water**, v. 13, n. 5, p. 705, 5 mar. 2021.

TURMAN, A. M., SOWBY, R. B., WILLIAMS, G. P., & HANSEN, N. C. Remote Sensing of Residential Landscape Irrigation in Weber County, Utah: Implications for Water Conservation, Image Analysis, and Drone Applications. **Sustainability**, v. 16, n. 21, p. 9356, 28 out. 2024.

VELOSO, M. F.; RODRIGUES, L. N.; FILHO, E. I. F. Evaluation of machine learning algorithms in the prediction of hydraulic conductivity and soil moisture at the Brazilian Savannah. **Geoderma Regional**, v. 30, p. e00569, set. 2022.

VON KROGH, G.; ROBERSON, Q.; GRUBER, M. Recognizing and Utilizing Novel Research Opportunities with Artificial Intelligence. **Academy of Management Journal**, v. 66, n. 2, p. 367–373, abr. 2023.

WANG, J., ZHEN, J., HU, W., CHEN, S., LIZAGA, I., ZERAATPISHEH, M., & YANG, X. Remote sensing of soil degradation: progress and perspective. **International Soil and Water Conservation Research**, 11(3), 429–454, 2023.

WANG, D., DONG, S., FU, S., SHEN, Y., ZENG, T., YU, W., LU, X., WANG, L., SONG, S., MA, J. Catalytic ozonation for imazapic degradation over kelp-derived biochar: Promotional role of N- and S-based active sites. **Science of The Total Environment**, v. 860, p. 160473, fev. 2023.

WEN, R.; LI, S. Spatial Decision Support Systems with Automated Machine Learning: A Review. **ISPRS International Journal of Geo-Information**, v. 12, n. 1, p. 12, 30 dez. 2022.

XIANG, Y., LIU, Y., NIAZI, N. K., BOLAN, N., ZHAO, L., ZHANG, S., XUE, J., YAO, B., & LI, Y. Biochar addition increased soil bacterial diversity and richness: Large-scale evidence of field experiments. **Science of The Total Environment**, v. 893, p. 164961, out. 2023.

XIE, G., ZHU, C., LI, C., FAN, Z., & WANG, B. Predicting the adsorption of ammonia nitrogen by biochar in water bodies using machine learning strategies: model optimization and analysis of key characteristic variables, **Environmental Research**, volume 267, 2025.

XIE, J., LATIF, J., YANG, K., WANG, Z., ZHU, L., YANG, H., QIN, J., NI, Z., JIA, H., XIN, W., & LI, X. A state-of-art review on the redox activity of persistent free radicals in biochar. **Water Research**, v. 255, p. 121516, maio 2024.

XIOURAS, C., CAMELI, F., QUILLO, G. L., KAVOUSANAKIS, M. E., VLACHOS, D. G., & STEFANIDIS, G. D. Applications of Artificial Intelligence and Machine Learning Algorithms to Crystallization. **Chemical Reviews**, v. 122, n. 15, p. 13006–13042, 10 ago. 2022.

XUE, H. et al. Runoff Estimation in the Upper Reaches of the Heihe River Using an LSTM Model with Remote Sensing Data. **Remote Sensing**, v. 14, n. 10, p. 2488, 23 maio 2022.

YAO, C., WANG, B., ZHANG, J., FAHEEM, M., FENG, Q., HASSAN, M., ZHANG, X., LEE, X., & WANG, S. Formation mechanisms and degradation methods of polycyclic aromatic hydrocarbons in biochar: A review. **Journal of Environmental Management**, v. 357, p. 120610, abr. 2024.

YAQUB, M., & LEE, W. Artificial intelligence models for predicting calcium and magnesium removal by polyfunctional ketone using ensemble machine learners, **Chemosphere**, volume 345, 140422, 2023.

YIN, R., LI, X., NING, Y., HU, Q., MAO, Y., ZHANG, X., & ZHANG, X. Machine learning unveils the role of biochar application in enhancing tea yield by mitigating soil acidification in tea plantations, **Science of the Total Environment**, volume 965, 2025.

ZHANG, T., TANG, Y., LI, H., HU, W., CHENG, J., & LEE, X. A bibliometric review of biochar for soil carbon sequestration and mitigation from 2001 to 2020. **Ecotoxicology and Environmental Safety**, v. 264, p. 115438, out. 2023.

ZHENG, H., CHEN, Q., TANG, B., ZHAO, M., LU, G., WANG, J., & SONG, G. Effectiveness evaluation method of marine environmental weapons and equipment based on ensemble learning. in: nakamatsu, k., kountchev, r., patnaik, s., abe, j.m., tyugashev, a. (eds) advanced intelligent technologies for industry. smart innovation, **Systems and Technologies**, vol 285. **springer, singapore**, 2025.

ZHOU, L. Perspective Chapter: Big Data and Deep Learning in Hydrological Modeling. Em: **River Basin Management - Challenges and Coping Strategies [Working Title]**. [s.l.] IntechOpen, 2024.

ZHU, H., AN, Q., SYAFIKA MOHD NASIR, A., BABIN, A., LUCERO SAUCEDO, S., VALLENAS, A., LI, L., BALDWIN, S. A., LAU, A., & BI, X. Emerging applications of biochar: a review on techno-environmental-economic aspects. **Bioresource Technology**, 2023.

APÊNDICE A – Artigo Publicado

GARCIA, F., ARAÚJO, M. S. B., DE SOUZA LEITE, A. C., CAVALCANTI GARCIA, R., GALVÍNCIO, J. D. Evolução da agricultura de precisão: uma revisão. **Revista Brasileira De Geografia Física**, v.17, n. 6, p.4761–4812, 2024. DOI: <https://doi.org/10.26848/rbgf.v17.6.p4761-4812>



Revista Brasileira de Geografia Física

Homepage: <https://periodicos.ufpe.br/revistas/rbgf>



Evolução da agricultura de precisão: Uma revisão

Fernando Antonio Marçal Garcia¹, Maria do Socorro Bezerra de Araújo², Antonio Celso de Sousa Leite³, Rodolfo Cavalcanti Garcia⁴, Josicleda Domiciano Galvêncio⁵

¹ Doutorando em Desenvolvimento e Meio Ambiente, Linha de Pesquisa Geoambiental, Rua Acadêmico Hélio Ramos, S/N, Cidade Universitária, 50740-530, Recife, PE, Brasil. (81) 21268275, fernandogarcia9999@gmail.com (autor correspondente). ² Professor Dr. Associado I, Universidade Federal de Pernambuco, Centro de Filosofia e Ciências Humanas, Departamento de Ciências Geográficas, Avenida Acadêmico Hélio Ramos s/n, Prédio do CFCH, 5o. andar, Sala 515, telefone: (81) 21268275, maria.baraujo@ufpe.br, Cidade Universitária, 50740530 - Recife, PE – Brasil. ³ Doutorando em Desenvolvimento e Meio Ambiente, Linha de Sensoriamento Remoto e Geoprocessamento, Professor EBTT, Instituto Federal do Piauí, Campus Teresina Central, Centro, 64000-040 – Teresina – PI - Brasil, e-mail: antonio.celso@ufpe.br. ⁴ Mestrando em Engenharia de Software, Linha de Pesquisa Sistemas Big Data, Arquiteto de Soluções, CI&T Company, Estrada Giuseppina Vianelle, 1455, Bloco C – Polo II - Alta Tecnologia, 13086-530 – Campinas – SP- Brasil, e-mail: rodolfovcgarcia@gmail.com. ⁵ Professor Dr. Titular, Universidade Federal de Pernambuco, Centro de Filosofia e Ciências Humanas, Departamento de Ciências Geográficas, Avenida Acadêmico Hélio Ramos s/n, Prédio do CFCH, 5o. andar, Sala 515, telefone: (81) 21268275, josicleda.galvencio@ufpe.br, Cidade Universitária, 50740530 - Recife, PE – Brasil.

Artigo recebido em 00/00/2024 e aceito em 00/11/2024

RESUMO

O crescimento populacional implica em demanda global por alimentos, impondo a busca de maior eficiência na produção agrícola mundial, diante da tendência de recursos naturais limitados e altos impactos ambientais. A agricultura de precisão vem evoluindo como uma solução promissora, ampliando o controle e monitoramento em tempo real das variáveis envolvidas no desenvolvimento eficiente e de qualidade do plantio, do transporte, estocagem e gestão de negócios, baseada em análises preditivas, suportadas por métodos estatísticos e técnicas de *business intelligence* e inteligência artificial. Estas ferramentas utilizam o tratamento destas informações para otimizar decisões e precificação. Entretanto, a adoção da agricultura de precisão ainda enfrenta barreiras significativas, como alto custo, falta de padronização de dados nos diversos dispositivos utilizados, complexidade técnica e necessidade de capacitação. Políticas públicas de fomento ainda são insuficientes, dificultando sua ampla utilização pelos produtores, principalmente em países em desenvolvimento. Novas tecnologias, como redes de comunicação mais seguras, padronização industrial, nanotecnologia, biotecnologia e algoritmos de aprendizado profundo de máquina têm aumentado a eficiência no mapeamento e inferência das variáveis agrícolas. Métodos como aplicação em taxa variável regulam uso de insumos e agroquímicos, contribuindo para minimizar impacto ambiental decorrente da atividade agrícola. Neste artigo uma análise abrangente da agricultura de precisão é realizada, analisando sua evolução, seus componentes usuais, atores participantes de seu desenvolvimento, políticas públicas aplicadas, benefícios e desafios enfrentados, auxiliando em sua compreensão e disseminação. Analisa tendências futuras e avanços em pesquisas, destacando seu potencial para influenciar práticas agrícolas mais sustentáveis, colaborando para minoração dos impactos ambientais decorrentes no avanço sobre biomas ainda preservados.

Evolution of precision agriculture: A review

ABSTRACT

Population growth implies a global demand for food, imposing the search for greater efficiency in global agricultural production, given the trend of limited natural resources and high environmental impacts. Precision agriculture has been evolving as a promising solution, expanding real-time control and monitoring of the variables involved in the efficient and high-quality development of planting, transportation, storage and business management, based on predictive analyses, supported by statistical methods and business intelligence and artificial intelligence techniques. These tools use the processing of this information to optimize decisions and pricing. However, the adoption of precision agriculture still faces significant barriers, such as high cost, lack of data standardization across the various devices used, technical complexity and the need for training. Public policies to promote it are still insufficient, hindering its widespread use by producers, especially in developing countries. New technologies, such as more secure communication networks, industrial standardization, nanotechnology, biotechnology and deep machine learning algorithms have increased the efficiency in mapping and inferring agricultural variables. Methods such as variable rate application regulate the use of inputs and agrochemicals, helping to minimize the environmental impact resulting from agricultural activity. This article provides a comprehensive analysis of precision agriculture, analyzing its evolution, its usual components, stakeholders involved in its development, public policies applied, benefits and challenges faced, helping in its understanding and dissemination. It analyzes future trends and advances in research, highlighting its potential to influence more sustainable agricultural practices, helping to reduce the environmental impacts resulting from the advance on biomes that are still preserved.

Introdução

Com o crescimento da população mundial nas últimas décadas, vários desafios associados precisam ser enfrentados, dentre eles, a crescente demanda mundial por alimentos (FAO-STAT, 2023; Khose e Mailapalli, 2024). Neste contexto, produtores mundiais de alimentos têm avaliado alternativas, disponibilizadas no mercado, para aumentar a eficiência da produtividade por hectare, para sustentabilidade da atividade. A agricultura de precisão, ou simplesmente AP (Gill e Chawla, 2021) adquiriu relevância por apresentar resultados na melhoria da gestão do negócio (Kim, 2023), monitoramento e otimização de recursos e variáveis envolvidas no processo (Chintakunta et al., 2023; Talaat, 2023), como condições do solo, água e clima, insumos e defensivos agrícolas, custos e previsões de produção, em todas as fases do plantio (colheita, transporte, armazenagem e comercialização). AP faz uso de modernas tecnologias industriais (Bazargani e Deemyad, 2024) e computacionais (Li et al., 2023; Saranya et al., 2023; Krishnababu et al., 2024) para compor sua solução, sem perder a possibilidade de integração com novos recursos e sistemas, inerentes ao negócio (Kumar e Bargavi, 2024; Maurya et al., 2024). Ela demonstra que sua importância para o setor está diretamente ligada à necessidade de modernização das áreas produtivas, especialmente em regiões onde os métodos tradicionais prevalecem, mas carecem de aumento da produção em escala, com minimização de custos e impactos ambientais (Rennings et al., 2024).

Em países desenvolvidos, a adoção de aplicativos de precisão para gerenciamento das culturas tem mostrado benefícios claros em termos de redução de custos operacionais, aumento da produtividade por hectare, além de mitigação de impactos ambientais negativos (Melzer et al., 2023). Entretanto, em países em desenvolvimento, a implementação da AP ocorre em percentuais menores, concentrada nos grandes produtores, com pequenos e médios produtores, em percentuais menores e variados, dependendo da região agrícola. Estes enfrentam desafios significativos, incluindo altos custos de aquisição, falta de infraestrutura tecnológica, baixa capacitação dos agricultores, adequação de políticas públicas de apoio e escassez de fomento para formação de cadeia de fornecedores (Huang, 2023).

Atualmente, o setor agrícola baseia-se em duas tendências: Competitividade, em termos de produtividade, com qualidade dos produtos a preços competitivos (Ezziyyani et al., 2023; Maurya et al., 2024); e eficiência, expressa na otimização dos custos de produção (Abobatta, 2021; Radhika e Masood, 2022), impactando na redução dos danos ambientais, premissa cada vez mais exigida pelos mercados consumidores como tendência em escala global.

De acordo com Kendall et al. (2022), uma das principais barreiras para a implementação da AP em larga escala é a falta de padronização das tecnologias disponíveis, o que dificulta a

integração de diferentes sistemas, aumentando os custos de implementação. Além disso, a carência de infraestrutura tecnológica, como conectividade de internet em áreas rurais, limita a eficiência das soluções baseadas em IoT, big data e aplicativos móveis, restringindo o potencial da AP em contextos de pequenos e médios produtores (Mitchell et al., 2021).

Diversos trabalhos publicados na literatura (Misara et al., 2022; Gawande et al., 2023; Khose e Mailapalli, 2024), tem ressaltado a importância de se avaliar os resultados obtidos com implementação de AP sobre os danos ambientais causados pela atividade agrícola. De maneira geral, o aumento da produtividade agrícola ainda ocorre pela ampliação das áreas de cultivo, impactando biomas diversos, com graves danos ambientais.

Embora as tecnologias de AP estejam avançando rapidamente, há uma lacuna significativa na adoção dessas soluções em nível global. Pequenos e médios agricultores, que representam a maioria dos produtores em muitas regiões, frequentemente não têm interesse na solução, facilidade de acesso ou capacidade financeira para implementar essas tecnologias, devido a barreiras de custos, falta de conhecimento técnico, falta de suporte local, políticas regulatórias e de financiamento inadequadas, dentre outras questões (Mizik, 2023; Nguyen et al., 2023). A centralização do desenvolvimento de AP em grandes produtores acentua as desigualdades no setor agrícola, limitando o potencial de transformação da AP como uma solução global.

Esse cenário tem estimulado os esforços alocados por diversas entidades, visando realização de novas pesquisas e desenvolvimento de bases tecnológicas, com pessoal especializado, visando fortalecer nos mercados produtores mundiais a ampliação da implementação da agricultura de precisão (Gawande et al., 2023).

A proposta central deste trabalho é disponibilizar uma fonte de consulta detalhada, propiciando o conhecimento aprofundado das diversas questões que envolvem a AP, por parte de pesquisadores, especialistas, profissionais e demais atores, envolvidos na disseminação, desenvolvimento e implementação da solução de AP, para que facilite a ampliação da discussão sobre o estado atual da solução nos diversos mercados usuários, demonstrando seu potencial, através dos benefícios disponibilizados, desafios

para tropicalização e integração com sistemas legados, as perspectivas futuras, com inovações tecnológicas que estão ou serão incorporadas. Além disso, fornece informações relevantes sobre fatores pouco ainda enfatizados em outros trabalhos de revisão, publicados na literatura, como a questão regulatória, as empresas fabricantes e os diversos tipos de aplicativos, a situação de políticas de fomento e financiamento, ressaltando seu papel para o crescimento da produção agrícola, contribuindo para minoração dos danos ambientais da atividade agrícola.

Este trabalho também procura estimular novas pesquisas para preencher lacunas na literatura, como a falta de estudos que abordem a real contribuição da AP, para a redução da necessidade de expansão de áreas agrícolas, ao tempo que procura responder, através dos diversos relatos e experimentos narrados nos artigos revisados, se realmente a AP poderá contribuir para aumento da produção mundial, minimizando impactos ambientais.

Por se tratar de uma solução que envolve integração de tecnologias de diversas áreas do conhecimento, a AP deve ser subsidiada de constantes publicações especializadas que contextualizem, de forma transparente e atualizada sua evolução, abstraindo a complexidade de seu desenvolvimento, melhorando a compreensão de seus componentes constituintes, as implicações legais para sua implantação e operação nos mais variados processos produtivos, além de clarificar os ganhos que podem ser obtidos pela otimização de produtividade por hectare plantado, a redução dos custos envolvidos, flexibilidade de roteio de culturas, e melhor gerenciamento de danos ambientais, aumentando a qualidade nutricional e segurança alimentar aos consumidores, com sustentabilidade ao empreendimento.

Outro foco importante deste estudo é estimular agtechs locais para o desenvolvimento de soluções modulares, de menor complexidade e custos, que sejam mais acessíveis e compreensíveis para pequenos e médios produtores. Essa abordagem modular pode permitir uma adaptação mais rápida e eficaz às diferentes condições regionais, facilitando a entrada da AP em nichos que hoje são plenamente atendidos pelas soluções tradicionais. Pequenos e médios produtores são responsáveis por uma parcela significativa das emissões de gases de efeito estufa, contaminação

por poluentes agrotóxicos, causadores de danos ambientais.

Por fim, o estudo visa demonstrar a necessidade de reformulação de políticas de financiamento e fomento direcionadas a mercados locais, por parte de autoridades e entidades formuladoras dessas estratégias, estimulando a criação ou fortalecimento de base tecnológica sólida e promissora, enaltecendo a importância da disseminação dos aplicativos de AP em todos os segmentos agrícolas produtivos, como estratégia de aumento da riqueza para países, uma ferramenta de preservação ambiental, contribuindo para a redução dos impactos negativos da agricultura nos ecossistemas envolvidos. A disseminação e evolução da AP dependem de um esforço conjunto de pesquisadores, agtechs (Singh e Kapoor, 2024), governos e produtores, de forma a transformar a agricultura em uma atividade mais eficiente, sustentável e inclusiva.

Material e métodos

Para elaboração deste artigo foi realizada ampla busca e análise de artigos de revisão da literatura, em artigos recentes ou similares, ou ainda, tratando sobre os elementos constituintes da solução de agricultura de precisão.

Coleta de dados

Os principais bancos de dados de produções científicas do Brasil e do exterior foram pesquisados, como em periódicos, Capes, Scielo, Google Acadêmico, Elsevier, IEEE – Xplore, Web of Science e Scopus, além de revistas especializadas em agricultura de precisão, agricultura digital. Também foram considerados tópicos relativos a instituições e entidades especializadas em políticas regulatórias e financiamento, fabricantes de componentes e empresas desenvolvedoras da solução de agricultura de precisão (AP), como agtechs (Brini, 2023) e fornecedores de serviços especializados, líderes no mercado, organizações agrícolas internacionais de fomento, tudo parametrizado no período dos últimos quatro anos, nos principais idiomas.

Desenvolvimento

Critérios de seleção das pesquisas

Como premissas da pesquisa, foram adotados como critérios de inclusão: (a) artigos revisados por pares, com bons graus de relevância, que discutiram tecnologias e aplicações de AP; (b) publicações de estudos de caso que explanassem sobre aquisição e implementação de AP, em grandes países produtores agrícolas; (c) pesquisas ou boletins econômicos relacionados ao tema, que versavam sobre os impactos econômicos e ambientais da AP.

Como critérios de exclusão foram considerados artigos que contivessem fontes de referências fora do parâmetro de buscas, ou promovessem abordagem superficial das soluções de AP.

Diversas expressões foram aplicadas nos atributos de busca das bases de pesquisas citadas: “digital farming”, “agricultura de precisão”, “agricultura 4.0”, “agtech”, “componentes AP”, “produtividade agrícola com agricultura de precisão”, “índices de vegetação”, “IoT na agricultura”, “sensores agrícolas”, “softwares agrícolas”, “soluções móveis agrícolas”, “inteligência artificial agrícola”, “sensoriamento remoto”, “drones agrícolas”, “computação embarcada na agricultura”, “regulamentação agrícola”, “redes sem fio na agricultura”, “financiamento agrícola”, “business intelligence na agricultura”, “big data na agricultura de precisão”, “riscos ambientais agrícolas”, “robótica na agricultura”, “computação distribuída na agricultura”, “blockchain na agricultura”, “selo verde”, “energia verde”, “danos ambientais na agricultura”, “tendências na agricultura”, dentre outras.

Procedimento de análise dados

O refinamento bibliográfico e busca avançada foi organizada utilizando-se a técnica PICO (Sousa et al., 2021). A ferramenta SciMAT (Cobo et al., 2012) foi aplicada para segmentar e classificar os arquivos selecionados pelos elementos da solução de AP. Por fim, as referências foram catalogadas com aplicativo Mendeley (2023). Foram gerados 278 artigos e 57 relatórios técnicos na sistemática acima. Ao final, após análise detalhada, 174 publicações foram selecionadas para produção desta revisão.

Nas aplicações de soluções de agricultura de precisão, são considerados pelas empresas que as elaboram, as chamadas agtechs (Tabela 1),

mecanismos e ferramentas estatísticas e computacionais, amplamente consolidadas no mercado. Fatores como complexidade de integração de diversos módulos constituintes, custos envolvidos, inovações decorrentes de pesquisa e inovação tecnológica (Figura 1). Levantamentos mostram que o mercado norte-americano investiu valores superiores a US\$ 2,36 bilhões, distribuídos em desenvolvimento e aperfeiçoamentos dos sistemas aplicativos (CEBDS, 2022), crescendo para US\$ 15,1 bilhões em 2020 (AgFunder,2022).

Os sistemas aplicativos de AP comumente ofertadas pelas diversas agtechs existentes no mercado contemplam módulos variados, como monitoramento remoto, controle de safras e riscos considerando clima, doenças, produção, armazenamento e logística, rentabilidade (colheita, projeção de preços, mercado, bolsa de valores).

A constante evolução deste tipo de solução é decorrente de uma variedade de avanços tecnológicos alcançados pela indústria. Entre estas inovações importantes estão o aperfeiçoamento do sistema de posicionamento global (GPS), WebGIS, sensoriamento remoto (SR), banco de dados relacionais para grande quantidade de dados (big data), equipamentos robóticos, drones (UAV), sensores inteligentes, técnicas de tratamento de taxas variáveis (VRT) com métodos de aprendizado de máquina (IA), dispositivos de IoT com sistemas embarcados, previsões e inferência de cenários para tomada de decisão com algoritmos de aprendizado profundo (DL), e mais recentemente, nanotecnologia e biotecnologia para melhoramento de insumos e agroquímicos utilizados.

Comparativamente, os resultados obtidos em produção agrícola utilizando AP é bastante significativo, em relação aos plantios tradicionais.

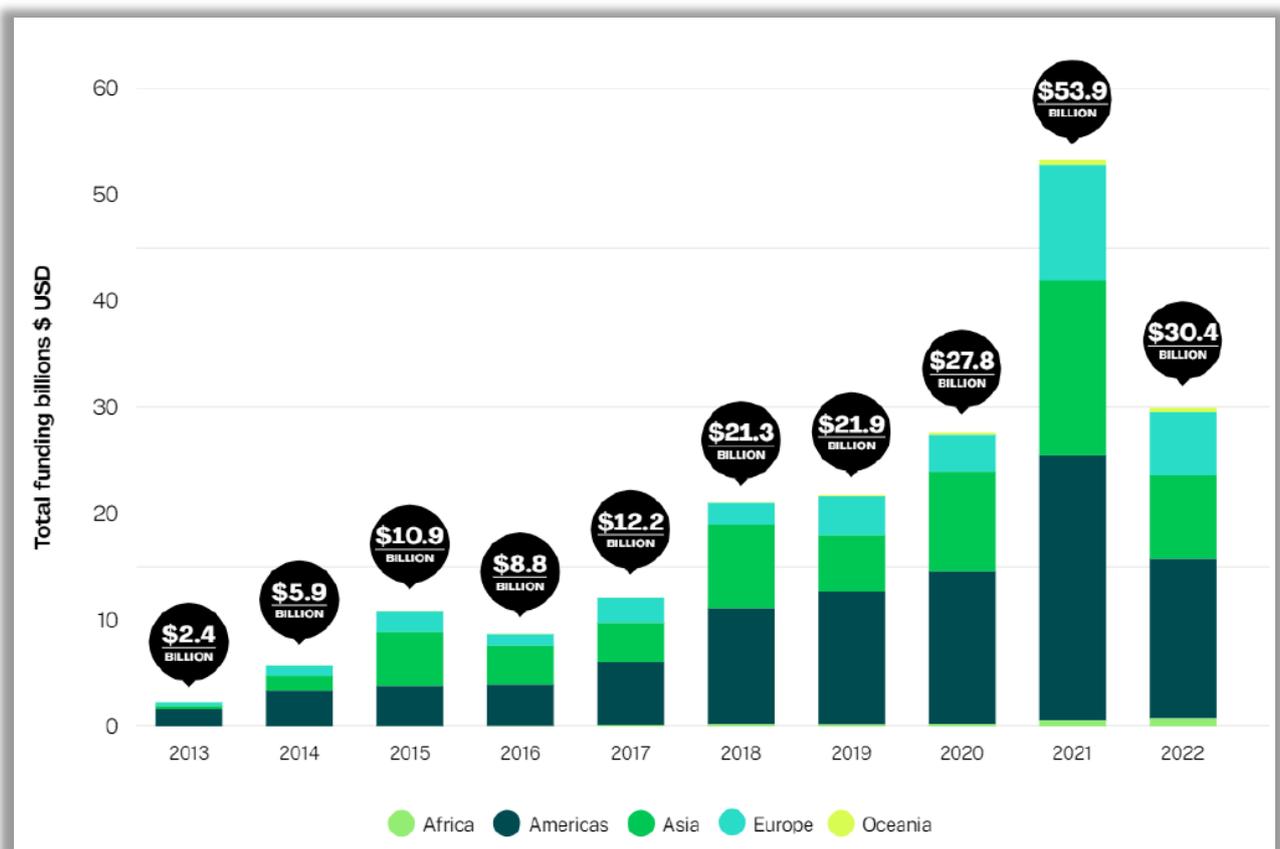


Figura 1: 10 anos de investimento em agtechs por região. Fonte: AgFunder (2022).

Dada a profusão e a complexidade das soluções tecnológicas existentes, esta revisão visa fornecer uma compreensão da estrutura básica da solução de agricultura digital, principais benefícios e desafios da agricultura de precisão com base em publicações recentes na literatura sobre o assunto.

Avanços tecnológicos com agricultura de precisão

O emprego de AP envolve o uso integrado de componentes apoiadas em tecnologia da informação e comunicação para coleta, tratamento e alocação de dados; automação robótica de equipamentos para operacionalização do plantio; tecnologia embarcada em sensores, satélites e drones, para tratamento, monitoramento e controle da produção. Estes sistemas aplicativos têm evoluído com incorporação de inovações tecnológicas (Sott et al., 2021), sendo comumente denominadas de AgroTIC. Em uma revisão bibliográfica realizada por Ragazou et al. (2022), são relatados diversos casos de aumento de produtividade com auxílio de ferramentas de agricultura de precisão. As soluções disponibilizadas geralmente são modulares (Tabela 1), com funcionalidades voltadas para *planejamento do plantio* (análise de solo, água, cultura, clima), *aplicação e monitoramento* (correção do solo, aplicação de insumos, controle de pragas, irrigação setorial, todos a taxas variáveis), *resultados* (previsão de colheita, estimativa de perdas, mapeamento de produtividade, avaliação de qualidade, estocagem), e *previsão de receitas* (controle de custos *online*, precificação e troca de produtos, em bolsas de produtos, como CBOT – Chicago/EUA, Rosário/Argentina e BM&F – São Paulo/Brasil). Em publicações na literatura, empresas como a Embrapa (2022) apresentam estudos de plantios realizados com AP, onde os resultados de produtividade da soja aumentaram em mais de 13%, atingindo 1,2 tonelada/hectare. Por fim, em outro estudo, da Universidade de Minnesota (U-Minn, 2022), plantações de milho atingiu e soja atingiram, respectivamente, 7,4% e 6,1% de produtividade de tonelada/hectare, utilizando aplicação de fertilizantes a taxa variável (VRT), comparativamente a outros plantios, com práticas agrícolas tradicionais na região.

Tabela 1: Agtechs de soluções de AP – Grandes produtores agrícolas. Fonte: Autores (2023).

Fabricante	País	Características	Funcionalidades ofertadas	Endereço eletrônico
Farm Works	USA	Gestão de plantações, gestão empresarial	Análise de mapas, dados de plantio/colheita, análise de dados coletados (solo, planta)	https://farmworks.com/
AgroSense	USA	Gestão de plantações, gestão empresarial	Análise e gestão de mapas, dados de sensores (solo, planta, plantas), transacionados por de big data e auxiliados por IA	https://www.agleader.com/
SMS	USA	Gestão de plantações, gestão empresarial	Mapas de análise, dados de sensores locais, análise de dados coletados (solo, planta)	https://www.agrosense.com/
AgDNA	USA	Gestão de plantações	Análise de mapas, dados de plantio/colheita	https://agdna.com/
Sentera	USA	Gestão de plantações	Análise e gestão de mapas, dados de sensores (solo, planta, plantas), transacionados por de big data e auxiliados por IA	https://sentera.com/
Farmers Edge	USA	Tomada de decisão, Geração, Plantio, Gestão empresarial	Gestão de insumos, análise de dados coletados (solo, planta), sistema de gestão agrícola de ponta a ponta	https://www.farmersedge.ca/
Climate Corporation	USA	Tomada de decisão, Geração, Plantio, Gestão empresarial	Análise de mapas, gestão de insumos, análise de dados coletados, sistema de gestão agrícola, precificação, créditos de carbono	https://climate.com/

Granular Corteva	USA	Tomada de decisão, Geração, Plantio, Gestão empresarial	Análise de mapas, gestão de insumos, análise de dados coletados, sistema de gestão agrícola, logística, precificação, créditos de carbono	<a href="https://granular.a
g/">https://granular.a g/
Indigo Ag	USA	Tomada de decisão, Geração, Plantio, Gestão empresarial	Análise de mapas, gestão de insumos, análise de dados coletados, sistema de gestão agrícola de ponta a ponta, créditos de carbono	<a href="https://www.indi
goag.com/">https://www.indi goag.com/
Conservis	USA	Monitoramento de plantio, sistema de gestão agrícola	Análise de mapas, gerenciamento de entradas, análise de dados coletados, sistema de gerenciamento agrícola de ponta a ponta	<a href="https://www.cons
erviscorp.com/">https://www.cons erviscorp.com/
Agworld	USA	Monitoramento de plantio, sistema de gestão agrícola	Análise de mapas, gestão de insumos, análise de dados coletados, sistema de gestão agrícola de ponta a ponta, módulo de negociação, análise de riscos ao meio ambiente e ecossistemas	<a href="https://www.agw
orld.com/">https://www.agw orld.com/
FarmLogs	USA	Monitoramento de plantio, sistema de gestão agrícola assistido por Inteligência Artificial, equipamentos automatizados	Análise de mapas, gerenciamento de entradas, análise de dados coletados, sistema de gerenciamento agrícola de ponta a ponta	<a href="https://m.farms.c
om/agriculture-
apps/crops/farml
ogs">https://m.farms.c om/agriculture- apps/crops/farml ogs
Raven	USA	Tomada de decisão, Plantio, Sistema de gestão empresarial assistido por Inteligência Artificial, Equipamentos automatizados	Análise de mapas, gestão de insumos, análise de dados coletados, sistema de gestão agrícola de ponta a ponta, módulo de negociação, análise de riscos ao meio ambiente	<a href="https://www.rave
nind.com/">https://www.rave nind.com/
The Yield	USA	Monitoramento de plantio, sistema de gestão agrícola	Monitoramento de plantio, sistema de gestão agrícola, terceirização de equipamentos	<a href="https://www.they
ield.com/">https://www.they ield.com/
Blue River Technology	USA	Monitoramento de plantio, sistema de gestão agrícola	Monitoramento de plantio, sistema de gestão agrícola por VRT, IA (Aprendizado de Máquina)	<a href="https://www.blue
rivertechnology.c
om/">https://www.blue rivertechnology.c om/
John Deere	Europa	Equipamentos automatizados, plantação, sistema de gestão agrícola	Equipamentos de campo automatizados, mapas de análise, dados de sensores locais, sistema de gestão agrícola	<a href="https://www.deer
e.com/">https://www.deer e.com/
Agrivi	Europa	Tomada de decisão, Geração, Plantio, Gestão empresarial	Gestão de insumos, análise de dados coletados (solo, planta, clima, doenças), sistema de gestão agrícola	<a href="https://www.agri
vi.com/">https://www.agri vi.com/
CropX	Europa	Monitoramento de plantio, gestão	Gestão de insumos, análise de dados coletados (solo, planta, clima), sistema de gestão agrícola de ponta a ponta	<a href="https://www.crop
x.com/">https://www.crop x.com/

Kubota	Europa	Tomada de decisão, Geração, Plantio, Gestão empresarial	Gestão de insumos, análise de dados coletados (solo, planta, clima), sistema de gestão agrícola de ponta a ponta	https://www.kubota-eu.com/
Farmdok	Europa	Monitoramento de plantio, gestão	Gestão de insumos, análise de dados coletados (solo, planta), sistema de gestão agrícola de ponta a ponta	https://www.farmdok.com/
Agroptima	Europa	Plantio, monitoramento de dispositivos	Gestão de insumos, análise de dados coletados (solo, planta), sistema de gestão agrícola de ponta a ponta	https://www.agroptima.com/
Drone Ag	Europa	Plantio, monitoramento de dispositivos	Gestão de insumos, análise de dados coletados (solo, planta) com drones, sistema de gestão agrícola de ponta a ponta	https://www.droneag.farm/
SenseFly	Europa	Plantio, monitoramento de dispositivos	Gestão de insumos, análise de dados coletados (solo, planta) com drones, análise de mapas, dados de plantio/colheita	https://www.sensefly.com/
Hórus Aeronaves	Brasil	Gestão de plantações	Mapas de análise, uso de drones, dados de sensores locais, dispositivos de monitoramento, previsão logística	https://horusaeronaves.com/
AgroSmart	Brasil	Gestão de plantações	Mapas de análise, uso de drones, dados de sensores locais, dispositivos de monitoramento	https://agrosmart.com.br/
Strider Syngenta	Brasil	Tomada de decisão, Geração, Plantio, Gestão empresarial	Gestão de insumos, análise de dados coletados (solo, planta), sistema de gestão agrícola de ponta a ponta, previsão de custos e vendas	https://www.strider.ag/
Solinftec	Brasil	Tomada de decisão, Geração, Plantio, Gestão empresarial	Monitoramento de plantio, gestão de insumos agrícolas, análise de dados coletados, sistema de gestão por IA, robótica, módulo de negociação	https://solinftec.com/
Smartbreeder	Brasil	Plantio, monitoramento de dispositivos	Análise de mapas, monitoramento de plantio, gerenciamento de insumos agrícolas, análise de sistema de gerenciamento de dados coletados por IA de previsão	https://www.smartbreeder.com.br/
Geoinova	Brasil	Gestão de plantações	Análise de mapas, monitoramento de plantio, gerenciamento de insumos agrícolas, análise de dados coletados, sistema de gerenciamento por IA	https://geoinova.com.br/
Agronow	Brasil	Geração de tomada de decisão	Mapas de análise, dados, inteligência artificial, custos e previsão de vendas	https://www.agronow.com.br/
Sintecsys	Brasil	Monitoramento/gerenciamento de incêndio	Mapas de análise, dados, câmeras, inteligência artificial, custos e previsão de vendas	https://brazillab.org.br/
Drakkar	Brasil	Tomada de decisão, Geração, Plantio, Gestão empresarial	Mapas de análise, dados, inteligência artificial, custos e previsão de vendas	https://www.drakkar.com.br/

Agrotools	Brasil	Plantio, empresarial	gestão	Mapas de análise, dados, inteligência artificial, custos e previsão de vendas	https://agrottools.com.br/
Agribolo	Índia	Monitoramento/gerenciamento de incêndio		Mapas de análise, dados, câmeras, inteligência artificial, custos e previsão de vendas	https://www.aibono.com/aibono
Crofarm	Índia	Tomada de decisão, Geração, Gestão empresarial	Plantio,	Mapas de análise, dados, câmeras, inteligência artificial, custos e previsão de vendas	https://www.crofarm.com/home
Beidahuang K. Seed	China	Plantio, empresarial, sementes de precisão, sistema de gestão assistido por Inteligência Artificial, equipamentos automatizados	gestão	Sistemas de plantio assistidos por GPS e VRT (<i>Variable Rate Technology</i>). Análise de dados de solo e clima para recomendações de cultivares.	https://beidahuan.com.cn
Meicai	China	Tomada de decisão, Geração, Gestão empresarial	Plantio,	Análise de mapas, gerenciamento de entradas, análise de dados coletados, sistema de gerenciamento agrícola de ponta a ponta	https://www.meicai.com/

A solução empregada na agricultura de precisão está relacionada ao tipo de cultivo, localização, condições ambientais, entre outros fatores. É importante ressaltar que práticas de agricultura de precisão muitas vezes exigem investimentos significativos em tecnologia e infraestrutura (Sanyaolu e Sadowski, 2024), o que pode aumentar os custos iniciais. Mas os benefícios de longo prazo da agricultura de precisão, incluindo aumento de produtividade e lucratividade, além da qualidade, muitas vezes superam esses custos (Hundal et al., 2023). A decisão do agricultor sobre o que controlar e gerenciar impacta na escolha da solução a ser adotada, tornando-se uma estratégia de negócios. Investimentos em pesquisa e inovação tecnológica implementados no desenvolvimento de soluções de AP direcionam os esforços das agtechs. A variada quantidade de aplicações voltadas para agricultura de precisão disponíveis no mercado, é resultante das diferentes funcionalidades requeridas pelo tipo e região de plantio, incluindo diferentes componentes tecnológicos (monitoramento remoto, controle de safra e riscos considerando clima, doenças, produção, armazenamento e logística) e de gestão do negócio (colheita, projeção de preço de venda e lucratividade, demanda mercado, câmbio).

A oferta de soluções de agricultura de precisão (AP), em franca atualização, está sendo vista por grande parte dos países produtores como

uma estratégia empresarial visando expandir o mercado, muitas vezes, associadas a ações promovidas por instituições de pesquisa e desenvolvimento, como associações técnicas (tabela 2).

Por exemplo, o mercado brasileiro inclui um segmento composto de empresas e órgãos governamentais promotores de ações voltadas à implementação de políticas públicas para fomento, regulamentação e amadurecimento de polo especialista na criação de tais soluções. Este mercado potencial inclui produtores, órgãos públicos, órgão de pesquisas e fabricantes industriais, para um público consumidor potencial em que somente 11% das terras produtivas utilizam soluções de AP (Embrapa, 2022; MAPA, 2022). Visando a promoção e desenvolvimento de aplicativos de AP, o país conta com diversas ações publicadas, apoiadas por diversos grupos de interesse nesse segmento de mercado, conforme demonstrado abaixo (Tabela 2). O ministério brasileiro para agricultura, pecuária e abastecimento (MAPA) tem sido um importante promotor e desenvolvimento de mercado consumidor de soluções de AP, apoiando a criação de grupo de trabalho intitulado “Câmara da Agricultura Digital”, que conta com diversas entidades associadas, dentre agências governamentais, empresas privadas fabricantes e prestadores de serviços, entidades de pesquisas agropecuárias. Para estimular o mercado

consumidor nacional a adotar as soluções de AP, o MAPA participou do lançamento do programa “Agro 4.0”, objetivando de estimular a implementação de tecnologias digitais, incluindo soluções de AP pelos agricultores brasileiros. Com este programa, o MAPA procura ofertar mecanismos de assistência financeira e técnica para seus associados. A Empresa Brasileira de Pesquisa Agropecuária (Embrapa) tem participado destas ações, subsidiando apoio técnico em pesquisas para o desenvolvimento de tecnologias de soluções AP, adaptadas à realidade agrícola brasileira, com dispositivos como sensores inteligentes, drones adaptados e sistemas modulares de suporte à decisão, testados e validados em diversas regiões produtoras do país.

Um dos maiores produtores mundiais de grãos, grande usuário das soluções de AP, o mercado norte-americano estimula, através da sua entidade governamental para agricultura, o Departamento de Agricultura, USDA (USDA, 2022) governamental, aquisição de soluções de AP por seus produtores nacionais, ofertada por imensa variedade de empresas agtechs. Para tanto, implementou várias políticas e iniciativas, reconhecidas pelo mercado produtor. O Serviço de

Conservação de Recursos Naturais (NRCS), vinculado ao USDA, oferta importante linha de crédito financeiro a agricultores que adotem tecnologias com soluções de AP, por meio do programa intitulado “Programa de Incentivo à Qualidade Ambiental – EQIP”. Já a Farm Service Agency (FSA), unidade governamental também vinculada ao USDA, há tempos oferta empréstimos financeiros, em condições diferenciadas, principalmente para pequenos e médios agricultores, estimulando-os a contratar serviços e tecnologias de AP. Além disso, o USDA promove intensamente parcerias com variadas entidades do setor privado, visando o desenvolvimento de pesquisas e formulação de novas tecnologias de AP. O USDA também tem se associado a empresas do polo tecnológico do Vale do Silício, realizando parcerias, com empresas como a *Blue River Technology*, com objetivo de desenvolver modernos sistemas de pulverização de agroquímicos por drones, auxiliados por dispositivos embarcados de visão computacional e aprendizado de máquina (ML), visando detecção de ervas daninhas em plantios, em tempo real.

Tabela 2: Principais entidades de fomento e regulação para soluções de AP. Fonte: Autores (2023).

Entidade	País	Propósito	Endereço eletrônico
ASBRAAP - Associação Brasileira de Agricultura de Precisão	Brasil	Associação empresarial sem fins lucrativos para o desenvolvimento científico e tecnológico, inovação e divulgação da aplicação de processos, técnicas e tecnologias da agricultura de precisão	https://asbraap.org/
CEBDS - Conselho Empresarial Brasileiro para o Desenvolvimento Sustentável	Brasil	Associação empresarial sem fins lucrativos que reúne os maiores produtores rurais do Brasil, promove o desenvolvimento sustentável, investe em pesquisa e apoia Agtechs na agricultura de precisão.	https://cebds.org/

CBAP - Comissão Brasileira de Agricultura de Precisão	Brasil	Grupo Executivo de Promoção de Ações de Fomento à Pesquisa em Agricultura de Precisão do Ministério da Agricultura do Governo Brasileiro.	https://www.gov.br/agricultura/pt-br
EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária	Brasil	Empresa brasileira ligada ao Ministério da Agricultura e Pecuária focada em pesquisa e desenvolvimento de soluções científicas.	https://www.embrapa.br/
Embrapii - Empresa Brasileira de Pesquisa e Inovação Industrial	Brasil	Promove a inovação na indústria brasileira, incluindo o setor agrícola, financiando projetos de pesquisa e desenvolvimento em parceria com empresas e instituições de pesquisa.	https://embrapii.org.br/
ABDI - Agência Brasileira de Desenvolvimento Industrial	Brasil	Entidade sem fins lucrativos que desenvolve projetos industriais brasileiros em parceria com os Ministérios da Agricultura, Economia, Ciência e Tecnologia e empresas de pesquisa industrial.	https://abdi.com.br/projetos/agro-4-0
Programa IA2MCTIC , do Ministério da Ciência, Tecnologia e Inovação MCTI	Brasil	Programa de apoio à pesquisa e projetos sobre o uso de inteligência artificial do Softex e do Ministério da Ciência, Tecnologia e Comunicações.	https://softex.br/ia/mcti/
Finep - Agência de Financiamento de Estudos e Projetos	Brasil	Agência governamental que financia a inovação e pesquisa científica e tecnológica no Brasil, com linhas específicas para agronegócio e tecnologias de agricultura de precisão.	http://www.finep.gov.br/
MAPA - Ministério da Agricultura, Pecuária e Abastecimento	Brasil	Órgão responsável por regulamentar políticas agrícolas, incluindo a utilização de tecnologias na agricultura de precisão.	https://www.gov.br/agricultura/pt-br
MCTI - Ministério da Ciência, Tecnologia e Inovação	Brasil	Responsável pelas políticas de ciência e tecnologia, incluindo regulamentações sobre inovação tecnológica no setor agrícola.	https://www.gov.br/mcti/pt-br
ANATEL - Agência Nacional de Telecomunicações	Brasil	Regulamenta o uso de radiofrequências e comunicações, relevantes para dispositivos IoT usados na agricultura de precisão.	https://www.gov.br/anatel/pt-br
SBA - Sociedade Brasileira de Agricultura	Brasil	Associação Brasileira de Automação - Comitê de Soluções em Agricultura de Precisão para o Agronegócio	https://sba.org.br/
INMET - Instituto Nacional de Meteorologia	Brasil	Agência do Ministério da Agricultura e Pecuária e provedora de informações meteorológicas.	https://portal.inmet.gov.br/
CNA - Instituto Nacional de Meteorologia	Brasil	Desenvolvimento de pesquisas, cursos e financiamento de soluções para agricultura, incluindo agricultura de precisão	https://www.cnabrazil.org.br/
ANA - Agência Nacional de Água e Saneamento Básico	Brasil	Gestão e monitoramento dos recursos hídricos nacionais, provedor de informações e fiscalização.	https://www.gov.br/ana/pt-br
Sebrae - Serviço Brasileiro de Apoio às Micro e Pequenas Empresas	Brasil	Oferece programas de suporte, consultoria e financiamento para micro e pequenas empresas, incluindo startups de agrotecnologia.	https://sebrae.com.br/sites/PortalSebrae

BNDES - Banco Nacional de Desenvolvimento Econômico e Social	Brasil	Oferece linhas de crédito e financiamento para projetos inovadores no setor agrícola, apoiando o desenvolvimento de tecnologias de agricultura de precisão.	https://www.bndes.gov.br/wps/portal/site/home
StartAgro - Instituição para fomento e financiamento de projetos de agricultura de precisão	Brasil	Plataforma que promove eventos, conteúdo e conexões entre startups do agronegócio, investidores e empresas, fomentando o ecossistema agtech no Brasil.	https://www.startagro.agr.br/
EIT Food - Instituição para a promoção e financiamento de projetos de agricultura de precisão	Europa	Comunidade europeia de inovação que apoia startups e projetos inovadores no setor agroalimentar, incluindo agricultura de precisão, por meio de programas de financiamento e aceleração.	https://www.eitfood.eu/
Horizon Europe - Instituição para a promoção de soluções de agricultura de precisão	Europa	O principal programa de financiamento de pesquisa e inovação da União Europeia, apoiando projetos em agricultura inteligente e tecnologias sustentáveis.	https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe_en
Agri-EPI Centre - Instituição para a promoção de soluções de agricultura de precisão	Reino Unido	Centro de inovação que colabora com empresas e instituições para promover tecnologias avançadas na agricultura, incluindo agricultura de precisão.	https://agri-epicentre.com/
Innovate UK - Empresa de desenvolvimento de soluções em agricultura de precisão	Reino Unido	Agência governamental britânica que financia projetos de inovação, incluindo soluções em agrotecnologia e agricultura de soluções.	https://www.ukri.org/councils/innovate-uk/
Farming-Futures, R&D Fund - Empresa de capital de risco	Reino Unido	Programa governamental que financia pesquisa e desenvolvimento em tecnologias agrícolas inovadoras.	https://iuk.ktn-uk.org/opportunities/farming-futures-rd-fund-climate-smart-farming/
USDA - Departamento de Agricultura do Governo dos Estados Unidos da América	USA	Desenvolvimento de política agrícola pública nos Estados Unidos da América. É ativo em pesquisa e ações para promover agricultura de precisão por meio de sua agência, NIFA.	https://www.usda.gov/
NIFA - Instituto Nacional de Alimentos e Agricultura do USDA	USA	Agência Federal para Alimentos e Agricultura. Realiza atividades de pesquisa, inspeção e promoção de tecnologia, incluindo AP	https://nifa.usda.gov/topics
FFAR - Fundação para Pesquisa em Alimentos e Agricultura	USA	Organização sem fins lucrativos que apoia pesquisas inovadoras em agricultura e alimentos, incluindo agricultura de precisão.	https://foundationfar.org/
AgFunder - Empresa de capital de risco, especializado em agtechs	USA	Plataforma de investimentos que conecta startups de agrotecnologia com investidores, fomentando o desenvolvimento de soluções inovadoras na agricultura.	https://agfunder.com/

The Yield Lab - Venture capital firm	USA	Aceleradora e fundo de investimento focado em startups de agtech, oferecendo capital e mentoria para desenvolver tecnologias agrícolas.	https://www.theyieldlab.com/
Plug and Play AgTech - Empresa de capital de risco	USA	Programa de aceleração que apoia startups de agtech com oportunidades de investimento, networking e colaboração com empresas do setor.	https://www.plugandplaytechcenter.com/industries/agtech
ISPA - Sociedade Internacional de Agricultura de Precisão	USA	Organização científica sem fins lucrativos que promove o desenvolvimento de pesquisas, congressos e publicações para avançar a agricultura de precisão em todo o mundo.	https://www.ispag.org/icpa
NAL - Biblioteca Agrícola Nacional do USDA	USA	Centro de Ciências Agrárias e Coleções de Pesquisa, incluindo Agricultura de Precisão do Governo Federal dos EUA.	https://www.nal.usda.gov/waic/precision-agriculture
PAFT - Força-Tarefa para Projetos de Agricultura de Precisão	USA	Seu objetivo é promover a adoção de técnicas de agricultura de precisão para melhorar a produtividade das colheitas, reduzir o desperdício e aumentar a sustentabilidade ambiental.	https://www.ams.usda.gov/precision-agriculture-task-force
AUVSI - Associação Internacional para Sistemas de Veículos Não Tripulados	USA	<i>Association for Uncrewed Vehicle Systems International</i> (AUVSI), a maior organização sem fins lucrativos do mundo dedicada ao avanço de soluções técnicas e robótica, representando corporações e profissionais de mais de 60 países envolvidos na indústria, governo e academia.	https://www.auvsi.org/who-auvsi
EPA - Agência de Proteção Ambiental dos Estados Unidos	USA	Agência promotora de políticas públicas para uso dos recursos ambientais (ar, solo, ar), monitoramento, fiscalização e provedora de informações. Inclusive para agricultura de precisão	https://www.epa.gov/
NSF - Fundação Nacional de Ciências	USA	Agência federal independente que promove diversas pesquisas, inclusive sobre agricultura de precisão.	https://www.nsf.gov/
FCC - Comissão Federal de Comunicações	USA	Regulamenta as comunicações e o uso do espectro eletromagnético, essencial para dispositivos conectados na agricultura de precisão.	https://www.fcc.gov/
ASABE - Sociedade Americana de Engenheiros Agrícolas e Biológicos	USA	Organização profissional que representa engenheiros, cientistas e outros profissionais que trabalham nas áreas de agricultura, alimentação e sistemas biológicos.	https://www.asabe.org/
Precision Ag Connectivity Task Force - Força-Tarefa para Revisão das Necessidades de Conectividade e Tecnologia da Agricultura de Precisão nos Estados Unidos	USA	Presidida pela FCC (Federal Communications Council) e envolvendo diversas agências federais com jurisdição sob a Lei Federal de Melhoria Agrícola (115-334, 132/2018) nos Estados Unidos da América, foi criada uma força-tarefa para revisar a necessidade de conectividade e tecnologia de agricultura de precisão no país.	https://www.fcc.gov/task-force-reviewing-connectivity-and-technology-needs-precision-agriculture-united-states

European Commission – DG de Agricultura e Desenvolvimento Rural	Europa	Formula políticas agrícolas da UE, incluindo regulamentos sobre tecnologia agrícola e agricultura de precisão.	https://commission.europa.eu/about-european-commission/departments-and-executive-agencies/agriculture-and-rural-development_en
EFSA - Autoridade Europeia para a Segurança Alimentar	Europa	Fornece aconselhamento científico e regulamenta aspectos relacionados à segurança alimentar, impactando as tecnologias de produção agrícola.	https://www.efsa.europa.eu/em
ENISA - Agência da União Europeia para Cibersegurança	Europa	Trabalha para garantir a segurança cibernética na UE, relevante para a proteção de dados em soluções de agricultura digital.	https://www.enisa.europa.eu/

Componentes da solução de AP

Embora a AP vise um sistema projetado para suportar o ciclo produtivo agrícola, poucas agtechs (Mendes et al., 2022, Singh e Kapoor, 2024) ofertam soluções modulares, contendo todas as fases do processo produtivo (Figura 2), especialmente no tratamento e monitoramento de variações espaciais e temporais nas condições de cultivo, como o estado da cultura, solo, plantas ou clima (Pasquel et al., 2022; Dos Santos et al.,

2023). No atual contexto, a maioria dos produtores ainda contratam solução de AP para atividades de preparo do solo, análise do solo, regulação e controle de equipamentos, no plantio e pulverização, com módulos contendo tratamento e análise de dados coletados. Módulos tratando de rotinas de apoio à decisão gerencial, logística, estocagem e precificação são menos contratados, embora em taxas crescentes (Brini, 2023). A integração desses componentes é fator chave para tornar a solução bem-sucedida.

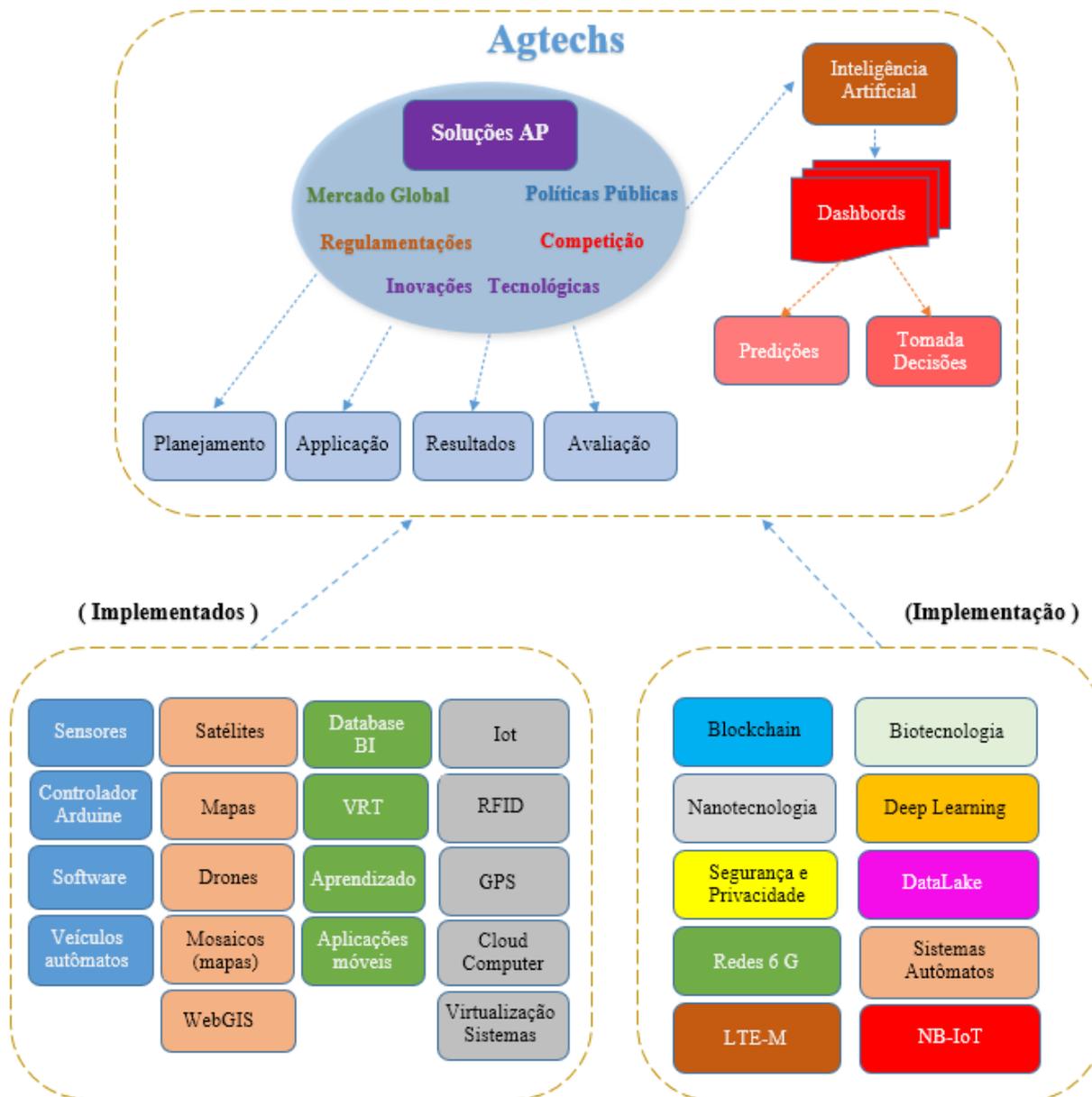


Figura 2: Diagrama de uma solução de agricultura de precisão. Fonte: Autores (2024).

Software de aplicação

Alguns fatores-chave no campo influenciam na decisão por adquirir um software aplicativo na solução de AP pelos produtores. Dentre estes, destacam-se a importância da análise das condições de pré-plantio, o monitoramento em quais fases do processo produtivo, além atividades de previsão de colheita, previsão para armazenamento e comercialização (Elahi et al., 2023; Maurya et al., 2024). Da análise realizada na literatura selecionada para este artigo, condições como a rapidez de atualizações e novas versões do aplicativo, novas funcionalidades disponibilizadas, a modularidade da solução, permitindo ao contratante montar o combo de sua aquisição personalizada, adaptáveis às necessidades de cada plantio, são diferenciais considerados na escolha do software aplicativo, e influenciam na estratégia planejada pelos produtores, como por exemplo, possibilidade de antecipar colheita (Figura 3). Dentre os módulos geralmente analisados pelos produtores, destacam-se: Processamento de informações coletadas em campo (Zhao et al., 2021; Qiao et al., 2022; Kganyago et al., 2024), processamento de imagens (Zhou et al., 2021; Huihui et al., 2023), processos de transferência de dados (Musa et al., 2023), geração de séries preditivas (Govindaraj et al., 2024), integração

com dispositivos móveis (Rivera et al., 2023), disponibilidade de informações na nuvem (Wei et al., 2023). Podemos considerar que, de forma geral, as soluções apresentam os seguintes módulos: (a) *coleta e análise de dados* (seleção de sementes, taxas de balanço de nutrientes necessárias do solo e das plantas, pragas, saturação hidráulica do solo, condições ambientais como temperaturas, previsão de chuvas, taxas de evapotranspiração), a partir de sensores, mapas por drones e/ou satélites, estações meteorológicas, dispositivos de campo automatizados, conectados por dispositivos de IoT, por exemplo; (b) *monitoramento do desenvolvimento do plantio*, fornecendo informações sobre diversas variáveis de forma *online* (taxas de irrigação, fertilização, pesticidas), geralmente tratadas e analisadas por aplicativos combinados de algoritmos de aprendizado de máquina, tecnologia de taxa variável, ou VRT, e plataforma de tecnologia GIS ou WebGIS, amparados em grandes volumes de dados armazenados em bancos de dados complexos, ou *big data*; (c) *painéis de modelos de ação ou dashboards* (previsões climática, pragas, saúde de solo e plantas, data colheita, produtividade, trocas produtos, geração créditos de carbono, estocagem, logística e precificação), alimentados em tempo real, com os relatórios gerados a partir das informações analisadas, servindo como ferramenta de suporte à decisão operacional, técnica ou de negócio.

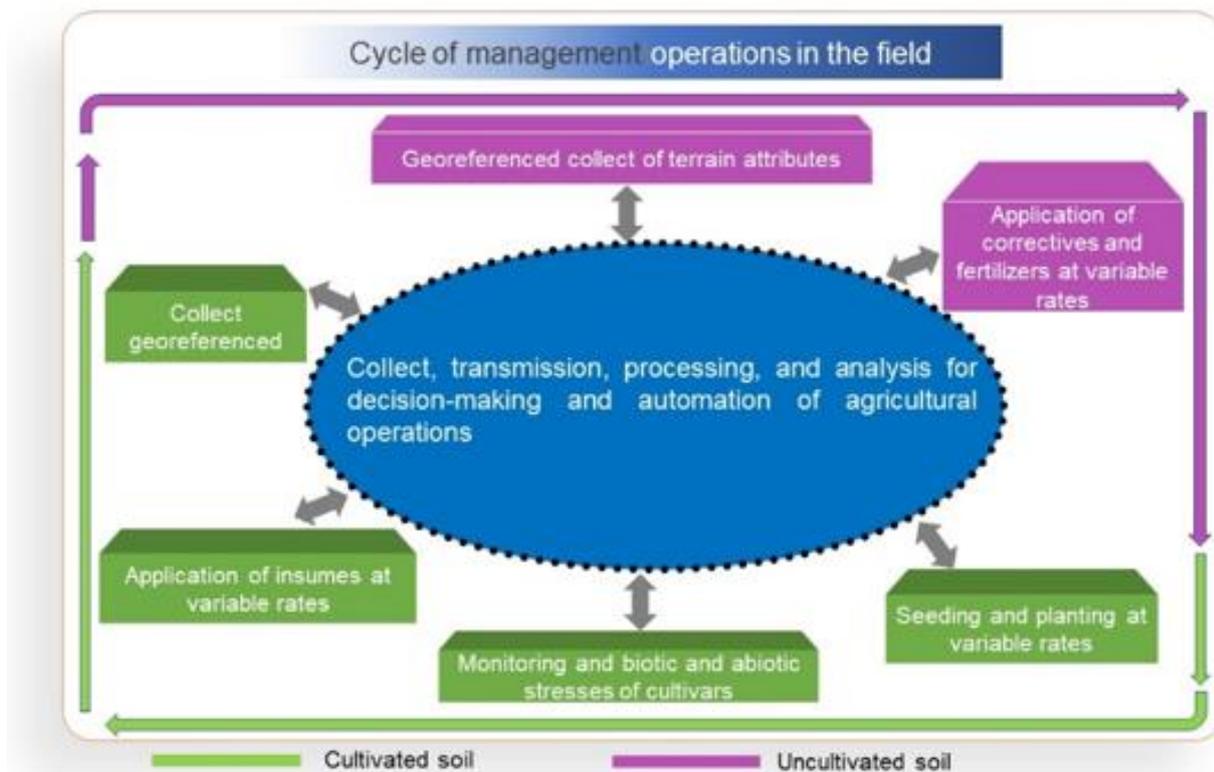


Figura 3: Funcionalidades do software PA. Fonte: Adaptado de Chaim *et al.*, (2014).

A composição dos custos de uma solução AP (hardware e software, instalação/configuração, capacitação e manutenção) ainda é alta, principalmente para mercados de países de baixa renda (Mizik, 2023; Rai e Kunte, 2024). Em geral, os módulos de coleta/processamento de dados e gestão da produção implicam em um custo percentual médio de 30 a 40% do valor total da solução (CEBDS, 2022; MAPA, 2022; Rimpika et al., 2023), sendo um dos motivos das diferentes

taxas de penetração da solução de AP em mercados produtores mundiais, como os norte-americanos, indiano, brasileiro e asiático. No mercado norte-americano, pequenos e médios produtores têm acesso a essas soluções com menores custos de aquisição, em torno de 20 a 25% (CEBDS, 2022; MAPA, 2022; Sanyaolu e Sadowski, 2024).

Outra funcionalidade importante oferecida por essas soluções é a geração e gerenciamento de enorme conjunto de dados sobre cada ciclo

produtivo, não apenas sobre as condições de cultivo e insumos utilizados, mas sobre todas as variáveis envolvidas nos plantios realizados, geralmente suportados por ferramentas de *machine learning* (Elahi et al., 2023; Aashu et al., 2024) e análise de dados com *business intelligence* (Raouhi et al., 2023). Essas séries históricas alimentam os módulos de predição e suporte à decisão.

Por fim, o software aplicativo é essencial na integração funcional entre os diversos dispositivos de hardware da solução (Raghunath e Devi, 2022; Mukati et al., 2024), como automação de máquinas, sensores e controladores de análise, que possuem software embarcado, em diversos padrões de fábrica.

Internet das Coisas (IoT)

A Internet das Coisas (IoT) é uma inovação tecnológica que possibilita a comunicação entre diferentes dispositivos instalados no campo e plataformas computacionais, de forma distribuída, em tempo real, desde que possuam o software adequado (Abu et al., 2022; Gonçalves et al., 2022; Hundal et al., 2023; Reda et al., 2023; NitinKamble e Kulkarni, 2024), facilitando a integridade das trocas de dados coletados (Tabela 3), amparados por disponibilidade de uma rede de comunicação no ambiente monitorado (Pasandideh et al., 2022). A correta classificação dos eventos coletados por dispositivos em campo, processados por meio do acesso a bancos de dados de informações, como é o caso de um sensor na identificação de uma praga (Micheni et al., 2022; Nyéki e Neményi, 2022), é um exemplo cotidiano da importância da tecnologia de IoT nas soluções de AP.

O aprimoramento da tecnologia de IoT está sendo padronizada em suas principais partes:

Dispositivos (hardware com sensores, atuadores, termostatos, *gateways*); *Conectividade* (padrões de comunicação como Wi-Fi, Bluetooth, Zigbee, RFID); *Processamento de dados* (em nuvem, plataformas de *edge computing* (He, 2024), com tratamento por algoritmos de IA); *Análise e disponibilidade de dados* (tratamento e geração de *insights*); e *Interface com usuário* (aplicativos Web e móveis, APIs para sistemas).

A integração de dispositivos com IoT de diversos equipamentos e componentes, como novos ambientes de inteligência artificial (Boopathi, 2024), computação em nuvem (Johnraja et al., 2024), bancos de dados interativos multimídia (Bhat e Huang, 2021; Sourav e Emanuel, 2021; Lu et al., 2022), processamento estratégico de informações (SIP), ou *business intelligence* com aprendizado de máquina profundo (Sanaeifar et al., 2023; Arya et al., 2024), é a espinha dorsal da popularização de uma solução típica de AP, criando uma poderosa ferramenta de processamento de informações para tomada de decisão, essenciais para sua disseminação no *agrobusiness*.

No entanto, ainda existem vários desafios associados à evolução da IoT na agricultura de precisão (Yazdinejad et al., 2021), principalmente nas questões de privacidade e segurança dos dados (Hundal et al., 2023), resultantes da coleta e transporte de dados obtidos das plantações, solo, clima, de tomada de decisão e sensíveis ao negócio, que devem ser protegidos contra ameaças cibernéticas, ou acessos não autorizados (Krishnababu et al., 2024; Ongadi, 2024). Apesar disso, novas pesquisas procuram disponibilizar protocolos de segurança de rede, criptografia e mecanismos de proteção nos dispositivos (Soussi et al., 2024), garantindo medidas aprimoradas para confiabilidade na solução de AP.

Tabela 3: Funcionalidades de IoT nos componentes das soluções de AP. Fonte: Autores (2023).

Fatores	Aplicação	Satélite	Sensores	UAV	Equipamentos Automatizados
Plantio	Eles monitoram a saúde das plantas e detectam pragas com insights abrangentes, complementados por informações detalhadas	Imagens (Landsat, Sentinel) para analisar a saúde das plantas (índices de vegetação e área foliar (IAF), temperatura do dossel, albedo da	Monitoramento do crescimento, biomassa e saúde das plantas usando sensores multiespectrais para	Imagens multiespectrais e hiperespectrais capturadas por drones para monitorament	Tratores e colheitadeiras com sensores integrados para medir o crescimento das culturas em tempo real e

coletadas que superfície, clorofila, o do ajustar os abrangem o dosagem de água), temperatura do crescimento e a identificação de dossel, índices de detecção de saúde das plantas. pragas e doenças. de vegetação, doenças, umidade foliar, pragas e saúde formato da do solo e das cultura, plantas. espessura do caule e altura da planta, nível d'água, transpiração.

Climáticos	<p>Fornecem informações detalhadas sobre o clima (precipitação, temperatura e umidade, velocidade e direção do vento, radiação solar, pressão barométrica, radiação solar, umidade relativa), sendo cruciais para decisões sobre irrigação, plantio e colheita.</p> <p>Dados climáticos de satélites (MODIS, Landsat) para análise de padrões climáticos e previsões meteorológicas em larga escala (radiação solar, velocidade e direção do vento, evapotranspiração, humidade, precipitação instantânea, fluxo de calor).</p>	<p>Medição de temperatura, umidade, velocidade do vento, radiação solar, evapotranspiração e outras ocorrências</p>	<p>Sensores climáticos embarcados para monitoramento de microclima local, gestão do plantio em tempo real.</p>	<p>Carregados com instruções de monitoramento, aplicam insumos, defensivos e outros agroquímicos, visando a evolução do plantio, com ajustes de irrigação, ventilação e controle de pragas e doenças</p>	
Solos	<p>Monitoram a umidade e a fertilidade do solo, permitindo uma gestão mais eficiente dos insumos agrícolas e a aplicação de nutrientes.</p>	<p>Análises de satélite para monitoramento de umidade do solo e detecção de variações na composição do solo.</p>	<p>Sensores de umidade, pH e condutividade elétrica para monitoramento em tempo real da condição do solo.</p>	<p>UAVs equipados com câmeras termográficas para monitorar a variação de temperatura do solo, umidade e transpiração.</p>	<p>Carregados com instruções de monitoramento, aplicam corretivos, fertilizantes, realizam análises de amostras de solo (pH, micro e macronutrientes minerais)</p>

Navegação	Dispositivos de GPS e RFID embarcados interoperam com os diversos componentes em campo, permitindo localização, correção e mudanças de rotas, rastreamento de insumos e coletas, propiciando troca de informações em tempo real.	Confirmação e/ou correção de dados satelitais auxiliados na navegação dispositivos de GPS, otimizando rotas de veículos autônomos e disposição de sensores em campo.	Dispositivos de GPS embarcados em tratores e veículos automatizados fornecem localização e dados em tempo real para otimização do trajeto.	Imageamento detalhado de variáveis monitoradas e geolocalizadas para otimizar trajetos e reduzir sobreposição de operações.	Localização, ajustes de rotas (dispositivos GPS), em veículos agrícolas autônomos, com roteamento programado, minimizando sobreposição, aumentando eficiência em aplicação insumos e agrotóxicos, e colheitas.
------------------	--	--	--	---	--

Big data

Grande quantidade de dados, coletados de diferentes dispositivos de IoT (Bramantoro et al., 2022; Alahmad et al., 2023) e de bases de dados externas, como USDA (USDA, 2022), requerem técnicas de padronização e armazenamento adequados de alta disponibilidade, visando suportar qualidade das informações processadas por poderosos algoritmos de aprendizado de máquina (Chintakunta et al., 2023), em técnicas de *business intelligence* (Bhat e Huang, 2021) e tecnologia de taxa variável (Ahmed et al., 2022; He, 2023; Krishnababu et al., 2024), constituindo assim, o principal papel das ferramentas de *big data*, que buscam garantir escalabilidade, disponibilidade e flexibilidade das transações para a solução de AP (Figura 4), em todas as suas etapas (Ding et al., 2023).

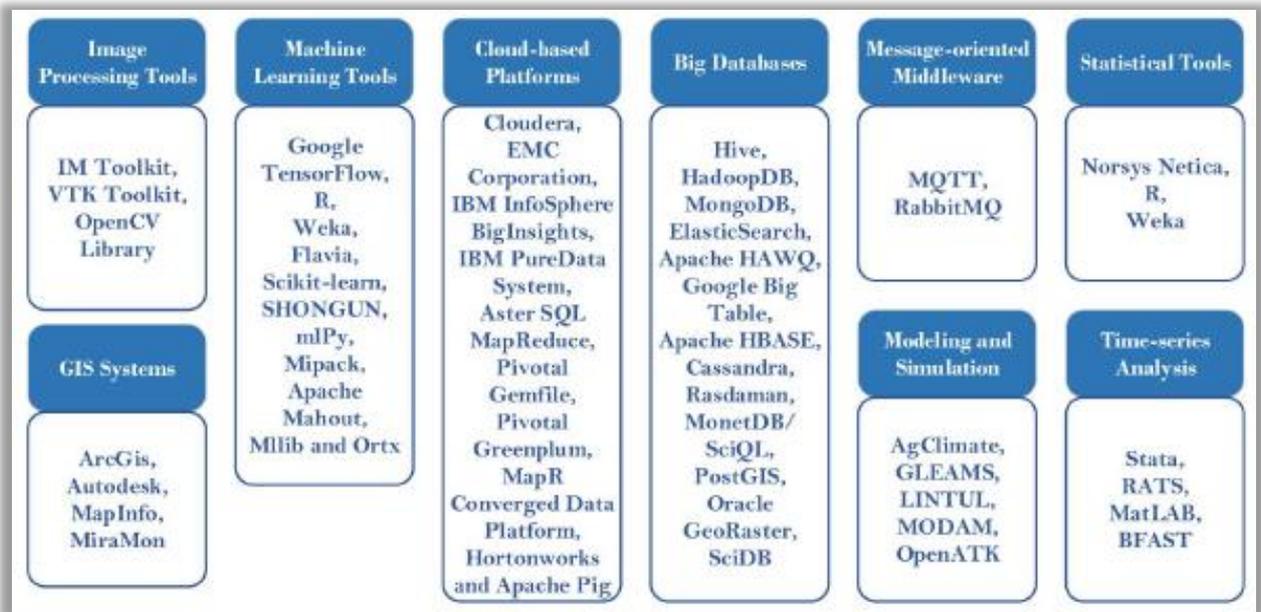


Figura 4: Ferramentas de software empregadas na análise de big data agrícola (Bhat e Huang, 2021).

Recentemente, com evolução da análise de variáveis espaciais em modelos de IA, os requerimentos de eficiência dos métodos analíticos de *big data* têm aumentado (Bhat e Huang, 2021), adicionando análises granulares das condições agronômicas (Dumont et al., 2023), facilitando a personalização das práticas agrícolas, de acordo

com as características específicas de cada área da propriedade, facilitando a tomada de decisões mais rápidas e precisas (Siddegowda e Devi, 2021; Elahi et al., 2023), otimizando não apenas a produtividade, mas também a gestão ambiental (Zaman, 2023).

Tecnologia de inteligência artificial

Considerando que a solução de AP envolve o domínio de diversas áreas do conhecimento (Cai et al., 2023; Chintakunta et al., 2023; Ijaz e Wozniak, 2024; Rai e Kunte, 2024), viabilizando a integração e o processamento de grande volume de dados, extraídos de diferentes fontes, com o objetivo de disponibilizar transações para

monitoramento, controle e previsão para tomada de decisões em tempo real, as inovações promovidas pela inteligência artificial estão gradualmente se tornando essenciais, diante das inovações trazidas, incorporando mais robustez e eficiência de previsões, inferências de cenários (Figura 5).

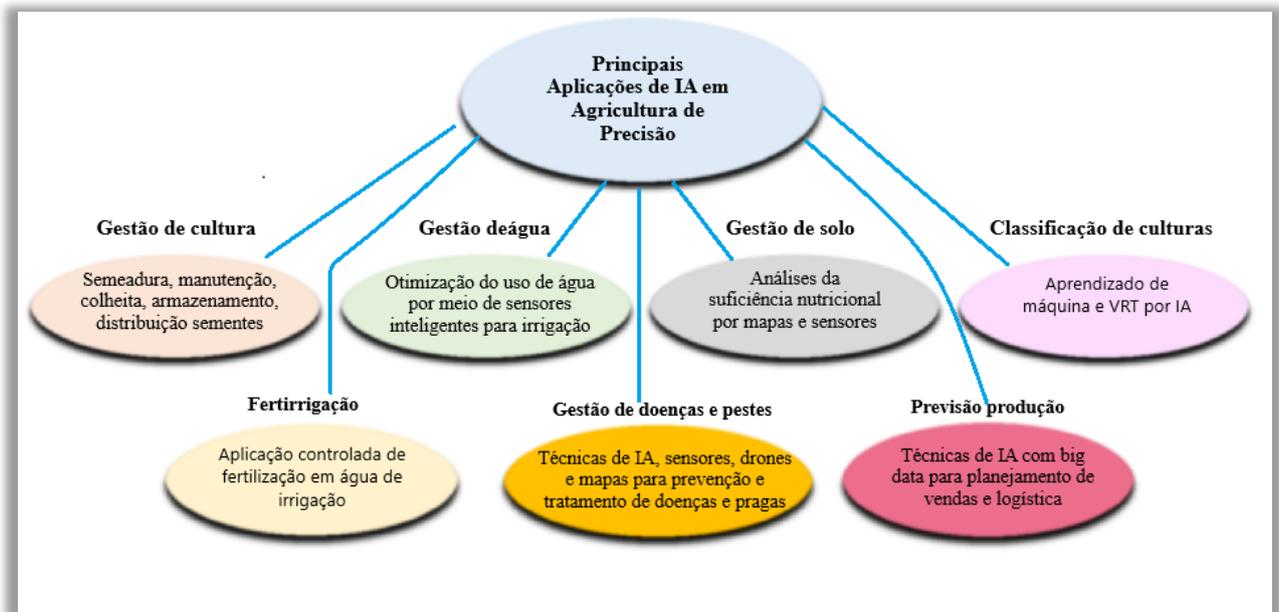


Figura 5: Principais aplicações de IA em solução de AP. Adaptado de Oliveira et al. (2023).

A implementação de ferramentas de inteligência artificial em soluções de AP contribui para melhorar a capacidade de processamento, eficiência, segurança e integração dos dados (Aashu et al., 2024; Sathya et al., 2024). Isso é potencializado pela evolução de dispositivos de campo e sistemas, como bancos de dados, sensores e IoT, pela incorporação de mecanismos como computação de borda (Zhang e Li, 2021; Akhtar et

al., 2021), processamento distribuído (Debauche et al., 2022; Fatima et al., 2024) e computação embarcada (Kumar et al., 2022), que permitem processamento de dados nativos diretamente destes componentes, em vez de depender de servidores centralizados da aplicação de AP, trazendo benefícios como redução de latência, economia de largura de banda e melhoria da segurança no tráfego de dados (Tabela 4).

Tabela 4: Mecanismos de Inteligência Artificial e seus impactos em soluções de AP. Fonte: Autores (2024).

Dispositivos AP	Fatores impactos	IA (modelo de aprendizado)	Algoritmos Específicos	Computação Embarcada	Computação de Borda	Processamento Distribuído	Federated Learning	Website
Sensores	Redução da latência e monitoramento em tempo real de variáveis do solo e clima	ML, DL	Redes Neurais Convolucionais (CNN), Regressão Linear, Árvores de Decisão	Sim	Sim	Sim	Sim	Alahmad et al., 2023, Soussi et al., 2024
Drones	Melhoria na coleta de dados, mapeamento detalhado, monitoramento de culturas	ML, DL	Redes Neurais Convolucionais (CNN), K-means, Support Vector Machines (SVM)	Sim	Sim	Sim	Sim	Sivakumar e Tyj, 2021; Govindaraj et al., 2024
Veículos automatizados	Maior eficiência operacional, redução de custos e otimização de rotas	ML, DL	Redes Neurais Recorrentes (RNN), Reinforcement Learning, Algoritmos Genéticos	Sim	Sim	Sim	Sim	Elahi et al., 2023; Sivasangari et al., 2023
Satélites	Monitoramento em larga escala, análise de clima e topografia	ML, DL	Árvores de Decisão, Random Forest, K-Nearest Neighbors (KNN)	Sim	Sim	Sim	Não	Misara et al., 2022; Fakhari e Khalid, 2023
Dispositivos IoT	Integração de dados em tempo real, otimização de insumos agrícolas	ML, DL	Redes Neurais Convolucionais (CNN), Aprendizado Preditivo, K-means	Sim	Sim	Sim	Sim	Saranya et al., 2023; Krishnababu et al., 2024
Equipamentos de irrigação	Otimização da distribuição de água, economia	ML, DL	Regressão Linear, Árvores de Decisão, Algoritmos de Reforço	Sim	Sim	Sim	Sim	Verma, 2021; Bramantoro et al., 2022

de
recursos

AD - Árvores de Decisão;

SVM - Support Vector Machines;

RNN - Redes Neurais Recorrentes;

RL – Aprendizado por Reforço;

AG - Algoritmos Genéticos;

RF - Random Forest;

KNN - K-Nearest Neighbors;

AP - Aprendizado Preditivo.

Tecnologia de taxa variável (VRT)

O primeiro componente estatístico de predição utilizado na solução de AP (Fakhar e Khalid, 2023; He, 2023; Szira et al., 2023), a tecnologia de taxa variável utiliza mapas gerados pelas soluções de GIS, contendo séries e dados estatísticos armazenados em big data, informações de variáveis diversas, coletadas em tempo real de sensores de campo, de sensoriamento remoto por drones e satélites, georreferenciadas por dispositivos IoT e GPS. Estes mapas de prescrições, como de variabilidade espacial (fertilidade do solo, saúde de plantas) são tratados por técnicas como o método de zonas de manejo, o método baseado em rendimento, o método de interpolação geoespacial, e o método de automação e controle (Ongadi, 2024).

A integração da técnica de taxa variável aos algoritmos de aprendizagem de máquina responde com melhor eficiência, com predições e inferências mais precisas, incorporando dados de séries temporais, informações frente a novos cenários inesperados, adequação a regulamentações públicas e de novas variáveis, por segmentos selecionados da lavoura, contribuindo para minimização de custos, qualidade e contaminação do ecossistema em seu entorno. Há relatos na literatura, como em Krishnababu et al. (2024), do aumento da produtividade de 20% e eficiência na minimização de custos com redução

do uso de agroquímicos e água, entre pequenos produtores na Índia, onde solução de AP com módulos de VRT foram implantadas. As agtechs têm na implementação da tecnologia de VRT como desafios atuais (He, 2023; Mizik, 2023): Os altos custos de implementação para pequenos e médios produtores; a complexidade na formatação de módulos com as técnicas de inteligência artificial; a interoperabilidade dos módulos de VRT com plataformas da solução de AP em sistemas legados, com diversos fabricantes de máquinas agrícolas, sensores, sistemas aplicativos diversos; a necessidade de pessoal especializado para implantação e manutenção do sistema, personalizados à realidade local. O relatório detalhado da *Straits Research* sobre o mercado de agricultura de precisão (Nair, 2023) aponta que, no mercado da América do Norte, as agtechs frequentemente ofertam as funcionalidades de VRT com geração de mapas de índices de variáveis de solo e plantas baseados em sensoriamento remoto por satélites (*satellite-based VRT*), ou por drones (*drone-based VRT*), onde os investimentos pesquisados junto a grandes clientes produtores, aponta prevalência atual para geração de mapas com drones.

Sensoriamento remoto (SR)

O principal objetivo do SR é a aquisição de informações sobre estado de variáveis, objeto do estudo da absorção e refletância da radiação eletromagnética por objetos na superfície da Terra (Bomoi et al., 2022; Kganyago et al., 2024; Manjunath et al., 2024; Messina e Modica, 2022; Omia et al., 2023; Pande e Moharir, 2023), caracterizado por aplicação de um conjunto de técnicas estatísticas e de reconhecimento de padrões, constituem o core do sensoriamento remoto. Este componente da solução de AP tem evoluído nos últimos anos, principalmente pela expansão das bases de satélites e otimização das técnicas de sensoriamento (Tabela 5), disponibilizadas nos últimos anos, permitindo a geração de informações de diferentes bandas espectrais (Radočaj, Plaščak, & Jurišić, 2023a;

Rivera et al., 2023). Recentemente, a associação com ferramentas de análise computacional (Abdelbaki e Udelhoven, 2022; Hashemi-Beni et al., 2022; Li et al., 2023), bem como com aplicação de nova camada nos mapas de prescrição de SIG, oriunda por imageamento por drones (Vishnu et al., 2024) tem otimizado os resultados de remote sensing. O SR é utilizado para gerar mapas de georreferenciamento da área escolhida para plantio, aplicando técnicas geoestatísticas para fornecer variáveis hiperespectrais e temporais, como índices de vegetação, temperatura, umidade, essenciais à análise das características do solo (Figura 6) e das plantas (Lara et al., 2021; Sun et al., 2022; Shukla et al., 2023), caracterizando em importante componente da solução de AP.

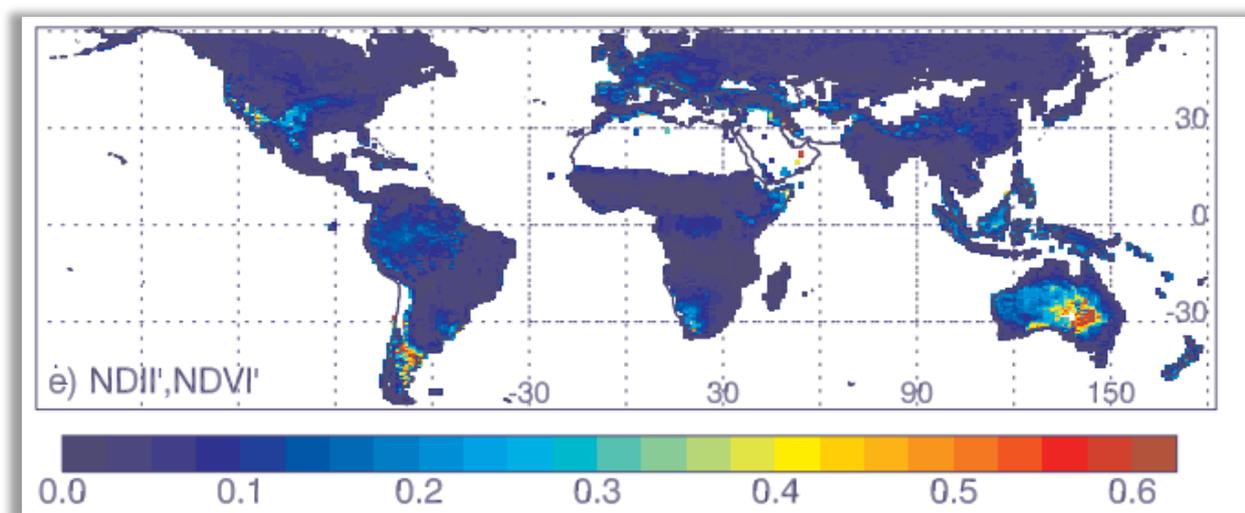


Figura 6: Mapa de condições do solo (Prey e Schmidhalter, 2019).

Tabela 5: Soluções de sensoriamento remoto (SR) para soluções de AP. **Fonte:** Autores (2023).

Plataforma Satelital	Funcionalidades Analíticas	Resolução Espacial	Frequência da Revisita	Aplicações	Endereço eletrônico
Landsat (NASA)	Monitoramento multiespectral (NDVI, estresse hídrico), detecção alterações no solo, mapeamento crescimento plantio	30 metros	16 dias	Análise saúde plantas, uso da água, mudanças solo e vegetação	https://landsat.gsfc.nasa.gov/
Sentinel-2 (ESA)	Imagens multiespectrais, monitoramento saúde da vegetação, detecção pragas e doenças, análise variabilidade do solo	10-60 metros	5 dias	Saúde plantas, análise variabilidade espacial, temporal e evapotranspiração, gestão de irrigação	https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2

MODIS (NASA)	Monitoramento climático, variabilidade plantas, solo e águas, detecção estresse hídrico. Mapeamento global	250 - 1.000 Metros	1 Dia	Visualização em larga escala saúde plantas, uso solo e recursos, variabilidade climática	https://modis.gsfc.nasa.gov/
SPOT (ESA)	Visualização detalhada crescimento plantas, uso solo, mapeamento terra e produtividade	1,5 - 6 Metros	2 - 3 Dias	Detecção alterações solo e vegetação, variabilidade climática, índices plantas	https://earth.esa.int/eogateway/missions/spot-1
Terra/Água (MODIS)	Análise temperatura superfície, monitoramento safras, variabilidade hídrica e solo	250 - 1.000 Metros	1 Dia	Análise saúde vegetação, alterações terra, índices vegetação, estresse climático	https://www.umb.edu/spectr/almass/terra-aqua-modis/

Drones

A popularização no campo pelos veículos aéreos não tripulados (vants), ou simplesmente, drones (Gokool et al., 2023; Rehman et al., 2023; Raouhi et al., 2023), incorpora vantagens nas soluções de AP, pois complementa, e em alguns casos, substitui o imageamento de imagens da área de plantio com qualidade de mosaicos gerados, além da análise e tratamento por aplicativos especializados, integrando com sistemas de informações geográficas. Os drones utilizados pela maioria das agtechs nas soluções de AP são de modelos de diversos fabricantes, dimensionados com peso médio de 1 kg, envergadura média de 1 m de diâmetro, possuindo software embarcado de navegação especializada de bordo, integrado com outros dispositivos de auxílio à navegação e transmissão de dados, como GPS, redes wifi, e dispositivos IoT (Tyagi e Pandey, 2024; Gamal et al., 2024), além de autonomia estimada de 1 hora de voo. Estes modelos geralmente são direcionados para o sensoriamento remoto ativo ou passivo, em geração de imagens, de acordo com a finalidade do imageamento (Fakhar e Khalid, 2023). Estes dispositivos móveis aéreos utilizam câmeras multiespectrais (para geração de índices como o

NDVI) e hiperespectrais (análises mais detalhadas da variabilidade da vegetação e solo, decorrentes da geração de índices como NDVI, NDWI, SAVI, EVI, SI, BSI, CI, WRI e RENDVI*), no sensoriamento remoto passivo da radiação refletida por plantas, solos e água (Radočaj, Šiljeg, Marinović, & Jurišić, 2023b; Govi et al., 2024). O SR realizado por drones diferenciam daquele gerado por emissões satelitais, importante para a solução de AP, devido a características destes equipamentos no imageamento (Figura 7), onde drones têm se tornado mais atrativos pelos custos decrescentes de sua aquisição, manutenção e flexibilidade na operacionalização de rotas, comparativamente aos satélites. Além disso, drones evoluem para incorporar outros dispositivos embarcados, permitindo integração com outros componentes da solução AP, como aplicativos produtores de mapas, sistemas SIG (Belcore et al., 2021; Das et al., 2023; Adewusi et al., 2024), além de renderização com inteligência artificial (Raghunath e Devi, 2022; Fakhar e Khalid, 2023), principalmente em modelos de aprendizado de máquina. Dependendo da finalidade planejada, podem trabalhar em substituição (Rejeb et al., 2022; Shahi et al., 2023), ou complementariedade ao imageamento por satélites (Tabela 6).

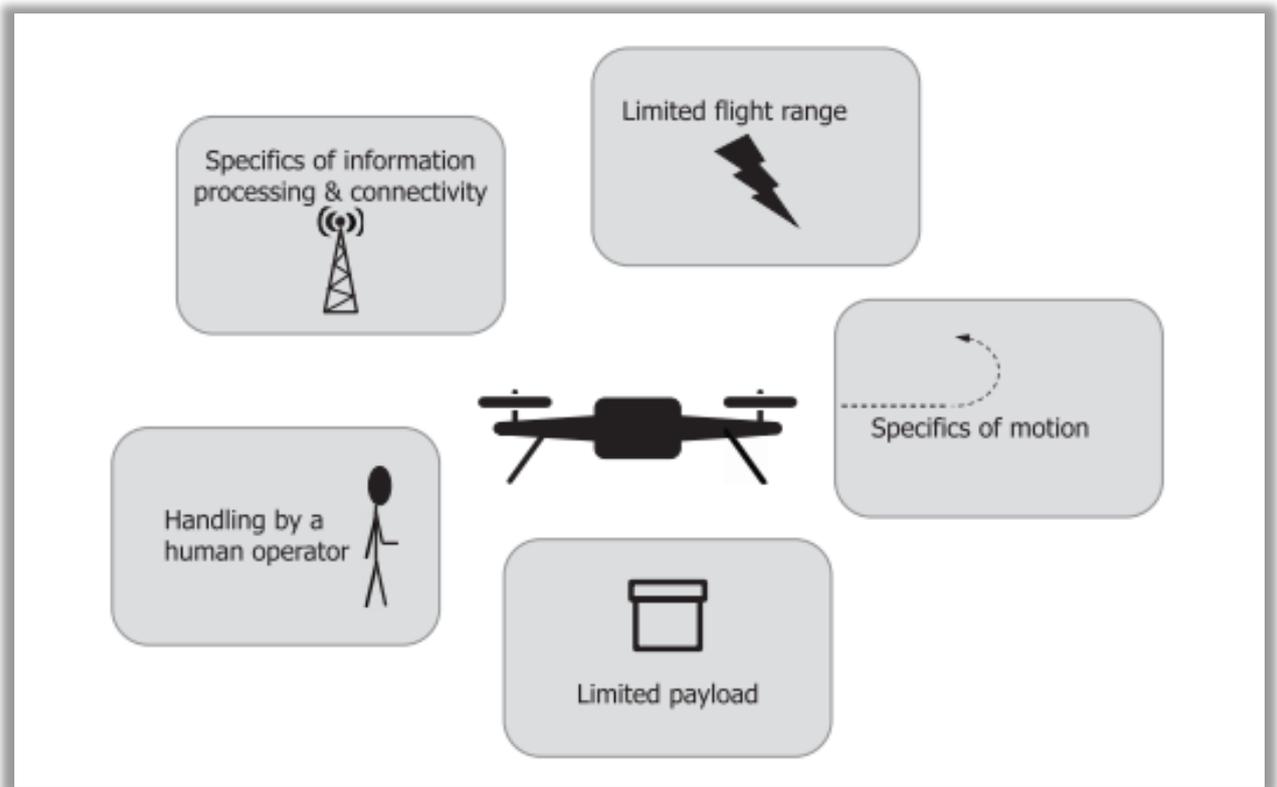


Figura 7: Esquema genérico de voo de drone (Otto et al., 2018).

Entretanto, a utilização de drones ainda enfrenta dificuldades, de acordo com a região, em questões ligadas à liberação regulatória burocrática; insuficiência de aplicações de hardware e software especializados para operação; requerer conhecimento especializado para processamento de imagens e integração de dados na solução AP; custo inicial total. Entretanto, pesquisas relatadas na literatura demonstram que seu uso vem crescendo e minimizando essas dificuldades, estimulado por inovações como uso de enxame de

drones; maior autonomia e automatização por navegação; minimização de falhas; prévia análise preditiva embarcada, otimizando os ortomosaicos; aumento da variedade de modelos por novos fabricantes; crescimento de fornecedores da cadeia de serviços; além de propostas de soluções de detecção de tentativas de intrusão por sistemas baseados em inteligência artificial (aprendizado de máquina) e tecnologia blockchain (Wei et al., 2023).

Tabela 6: Sensoriamento remoto por Drones versus Satélites para soluções de AP. Fonte: Autores (2023).

Características técnicas	Drones	Satélites
Resolução espacial	Alta	Moderada
Cobertura espacial	Limitada (autonomia Drone)	Ampla (grandes áreas)
Frequência de monitoramento	Alta (sob demanda)	Moderada/baixa (órbitas fixas)
Custo operacional	Moderado (de acordo com funcionalidades/manutenção)	Variável (de acordo com funcionalidades e resolução dados analisados)

Condições meteorológicas	Teto operacional de acordo com ventos e chuvas	Teto operacional de acordo com bloqueio nuvens (de radar especialmente)
Treinamento necessário	Navegação, aplicativos, tratamento de dados	Aplicativos, tratamento dados
Funcionalidade mais utilizadas	<ul style="list-style-type: none"> * Monitoramento em alta resolução de pequenas áreas agrícolas; * Aplicação de insumos como fertilizantes e pesticidas com alta precisão; * Monitoramento de pragas, doenças e saúde das plantas com câmeras multiespectrais; * Criação de mapas NDVI (Normalized Difference Vegetation Index) para avaliar a saúde da vegetação. 	<ul style="list-style-type: none"> * Monitoramento de grandes áreas agrícolas de uma só vez; * Detecção de variabilidade climática e mudanças em larga escala (como seca ou enchentes); * Uso de imagens multiespectrais e hiperespectrais para análise de saúde das plantas e produtividade; * Análise de índices espectrais, como humidade, temperatura e evapotranspiração.
Aplicativos usuários	<ul style="list-style-type: none"> * <u>DroneDeploy</u>: Mapeamento de campo em tempo real, análise de produtividade e planejamento espaço; * <u>Pix4D</u>: Criação de mapas geoespaciais de alta precisão; * <u>DJI Terra</u>: Mapeamento para planejamento e análise agrícola com drones DJI; * <u>PrecisionMapper</u>: Análise de safras, criação de mapas NDVI e outros índices; * <u>QGIS</u>: Processamento de dados e dados multiespectrais, integrando-os com outras camadas de dados geoespaciais; * <u>EEFlux</u>: Processamento dados com Landsat na estimação índices de evapotranspiração e gestão hídrica; * <u>SUPER</u>: Plataforma de monitoramento de variáveis vegetais, como biomassa, nutrição e crescimento plantas; 	<ul style="list-style-type: none"> * <u>Google Earth Engine</u>: Monitoramento global com divers indicadores para análise agrícola; * <u>FieldView</u> (da Climate Corporation): Processamento de dados para monitoramento em tempo real de diversas variáveis nas plantações; * <u>Sentinel Hub</u>: Processamento de dados com uso dos satélites Sentinel e Landsat, fornecendo insights sobre saúde das plantas; * <u>Cropio</u>: Plataforma de gerenciamento agrícola que gera painéis para tomada de decisão no campo; * <u>QGIS</u>: Processamento e visualização de dados com imagens Landsat e Sentinel, para geração de análise geoespacial, criação de mapas NDVI, monitoramento da saúde das culturas e recursos hídricos. * <u>Google Earth Engine</u>: Dados de evapotranspiração e gerenciamento de água, calculados usando o método METIC, disponibilizados. * <u>Landsat 8</u>: Modelagem de variáveis hidrológicas, evapotranspiração e qualidade da água, modeladas com SWAT, permitindo integração com outras ferramentas de SR, como drones.

	* Power (NASA): Plataforma da NASA para processamento dados para geração de variáveis meteorológicas, como radiação solar, temperatura, precipitação;	* CERES/MERRA-2/GISS: Geração de modelos de radiação solar e evapotranspiração, analisados pelo MERRA-2, e capazes de integrar com modelos climáticos do GISS. Geralmente usa o método do Modelo Penman-Monteith.
Regulamentação (USA)	Legislação regulamentada pela FAA (Federal Aviation Administration, com atenção a Parte 107 para atividades comerciais (altura, locais), exigindo licenciamento de pilotos, além de restrições em áreas próximas a aeroportos.	Legislação normatizada pela NOAA (National Oceanic and Atmospheric Administration) e o Departamento de Comércio dos EUA, com especificações detalhadas para privacidade e segurança nacional, principalmente para compartilhamento de dados e imagens no território dos EUA e exterior.
Regulamentação (Brasil)	Legislação regulamentada pela ANAC (Agência Nacional de Aviação Civil), inclusive com registro no sistema SISANT (altura voos, locais), com registro voo no DECEA (Departamento de Controle do Espaço Aéreo) a áreas próximas a aeroportos.	Legislação supervisionada pelo INPE (Instituto Nacional de Pesquisas Espaciais), inclusive com direito de uso de imagens e dados geoespaciais para atividades comerciais sujeitos a leis como lei de proteção de dados e privacidade (LGPD), além de acordos internacionais.

(*) NDVI - Avaliar a saúde da vegetação;
 NDWI - Medição de água nas plantas;
 PRI – Fotossíntese e metabolismo plantas;
 SAVI – Monitorar vegetação em solo exposto;
 EVI – Avaliar áreas de vegetação densa;

SI – Avaliar fertilidade e estrutura do solo;
 BSI – Avaliar qualidade solo/áreas de solo exposto;
 CI – Avaliar correções de fertilidade do solo;
 WRI – Monitorar áreas irrigadas/acúmulo de água;
 RENDVI – Detecção estresse hídrico (pragas);

Sistemas de Informações Geográficas (WebSIG)

Os chamados sistemas de informação georreferenciados, ou simplesmente WebSIG, são frequentemente usados em aplicações de AP (Pasquel et al., 2022; Hu, 2023; Karada et al., 2023) para fornecer informações precisas sobre variações geoespaciais georreferenciadas, como mapas incluindo imagens de solo, planejamento e dados meteorológicos, resultando em visualização de padrões e tendências em diversas camadas dinamicamente funcionais, (Figura 8), auxiliando a tomada de decisão (Kurniawan et al., 2023). A geração de mapas com nível de precisão está sendo disponibilizada em servidores remotos (computação em nuvem) ou locais, através de sistemas WebGIS, estas informações integradas

com outras funcionalidades da solução AP, como por exemplo, a disponibilização de produtos e serviços de sensoriamento remoto com WebGIS por agências norte-americanas (NASA, 2022; USDA, 2022). Este componente da solução AP está associado ao uso conjunto com ferramentas de GPS, sensoriamento remoto (SR), mosaico de mapas por drones, conexão de IoT, Internet, entre outros.

Os benefícios alcançados com sua implementação, como auxílio na tomada de decisão, com informações em tempo real, permitem analisar e decidir sobre diversas variáveis específicas e sensíveis, permitir o monitoramento da produção, condições ambientais e cultivo. As soluções desenvolvidas pelas agtechs

buscam amenizar problemas ainda presentes em sua adoção, como a integração de dados com diferentes formatos e padrões de tratamento de diferentes fontes, questões relacionadas à privacidade e segurança, necessidade de conhecimento técnico para operar essa ferramenta, bem como garantir a conectividade dos diversos

elementos constituintes da geração de informações. As novas versões de aplicações de agricultura de precisão já oferecem alternativas como o processamento e a disponibilização das informações do SIG em uma nuvem de dados (Tabela 7).

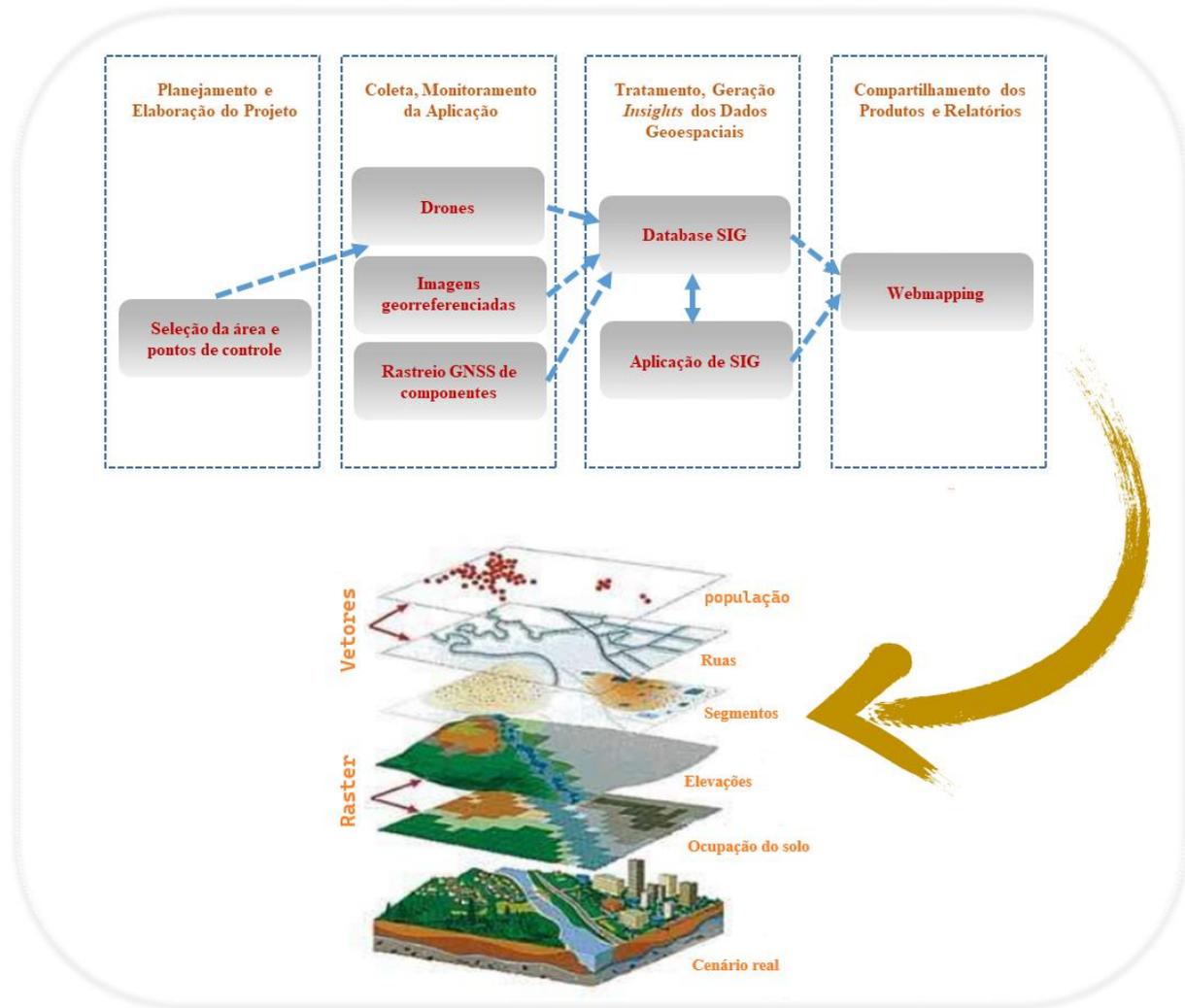


Figura 8: Composição de um SIG em camadas. Fonte: Adaptado de Rocha *et al.* (2015).

Tabela 7: Softwares aplicativos de SIG na agricultura de precisão. Fonte: Autores (2023).

Funcionalidades da Plataforma	GIS	WebGIS
Acesso	Em computadores locais ou redes internas, com exigência de licenças ArcGIS ou QGIS	Aplicação em nuvem ou servidor em rede, com acesso via Web, sem necessidade de software local
Infraestrutura mínima	Requer estações robustas, com alta capacidade de processamento e armazenamento.	Requer estrutura local menos robusta, com processamento e armazenamento remoto

Processamento de dados	Grande volume de dados de forma local, em análises geoespaciais, como modelagem 3D	Visualização local do resultado do processamento em servidores remotos, com limitação para análises complexas
Colaboração	Limitada, dados necessitam de compartilhamento entre os diversos usuários da aplicação	Compartilhado, acesso distribuído, em tempo real, com acesso simultâneo de edição de dados por múltiplos usuários da aplicação
Atualização de dados	Sincronização local, e caso dados estejam alocados em máquinas diferentes, requer alto controle de versionamento de dados	Sincronização em tempo real para qualquer atualização realizada pelos usuários da aplicação, garantindo controle de versionamento de dados
Custo estimado	Requer licenças de software aplicativo por estação de uso, considerando ainda, os requisitos mínimos por máquina instalada	Requer licenças de software por servidor, considerando o número de usuários, com requisitos de máquinas locais usuárias menores. Há o custo adicional de assinatura para serviços e hospedagem, caso seja em nuvem.
Interatividade	Permite personalização de análises, relatórios e visões espaciais.	Customizado para personalização de dados, relatórios e visões espaciais, mas excelente poderoso para interações básicas das aplicações compartilhadas
Conectividade	Acesso local, sem exigência de saídas para a Internet	Exigência de saída Web para acesso de ferramentas e dados, principalmente para alocação em nuvem.
Manutenção	Controle local de atualizações de software aplicativo e hardware	Responsabilidade remota no servidor de aplicações por atualizações centralizadas
Segurança de dados	Maior controle por armazenamento local, mas exige cuidado com backups e protocolos de segurança internos	Segurança sob responsabilidade em quem aloca os servidores remotos, ou aplicação em nuvem, exigindo confiança em provedores quanto à integridade dos dados
Utilização na AP	Utilizados por Agtechs nas versões mais antigas de suas soluções, ou por exigência personalizada de clientes, sendo ainda excelente para criação de mapas detalhados de solo, culturas e clima, com análises profundas e personalizadas, principalmente em locais com deficiência de cobertura de redes de Internet	Maioria das soluções ofertadas por Agtechs, principalmente nos mercados com maior infraestrutura e cobertura de redes de Internet, sendo excelente opção para compartilhamento para usuários remotos em qualquer lugar, inclusive, via rede móveis, facilitando a colaboração

Sensores

Esses dispositivos promovem a coleta em tempo real das principais variáveis envolvidas na produção agrícola, como das condições do solo, plantas e clima (Shukla et al, 2023; Musa et al., 2023), através de drones, satélites e dispositivos de IoT, justificando o uso desses mecanismos em campo, por soluções AP (Alahmad et al., 2023; Karunathilake et al., 2023; NitinKamble e Kulkarni. 2024).

Sensores podem trabalhar em conjunto com os atuadores, desempenhando importante papel na coleta e controle de atividades corretivas de variáveis essenciais (Manjunath et al., 2024), por setores do plantio, como umidade do solo, temperatura, pH e situação de nutrientes (Peppi et al., 2023; Soussi et al., 2024), detecção de eventos como pragas (Qiao et a., 2022). Por exemplo, os atuadores controlam dosagem nos sistemas de irrigação, operando de acordo com parâmetros de ajustados por taxas variáveis de aplicação de fertilizantes, ou acionar pulverizadores para

aplicação de pesticidas ou herbicidas, ou ainda, para controle de aspersores de água.

As últimas gerações de sensores utilizados nas soluções de AP, estão sendo empregados com: Sistemas baseados em Arduino com software embarcado, controlados manualmente, ou remotamente; associados com tecnologias VRT, ajustando a dosagem de insumos (Mizik, 2023); com algoritmos de aprendizado de máquina (Li et al., 2023; Omia et al., 2023), para processamento e geração preditiva de informações coletadas, integradas com dispositivos de IoT (Das et al., 2023; Shahi et al., 2023), e big data (Bhat e Huang, 2021; Nyéki e Neményi, 2022), para reconhecimento de padrões e otimização de tomadas de decisões no campo.

De acordo com as funcionalidades e parâmetros requeridos pela aplicação do AP, os sensores podem ser térmicos, ópticos (Pande e Moharir, 2023) ou elétricos (Kim e Lee, 2022), geralmente transmitindo seus dados para um computador, que

processa as informações coletadas via rede sem fio (Huerta et al., 2021; Ragazou et al., 2022). Eles também podem ser remotos, como as características de sensoriamento remoto, ou sem fios (Musa et al., 2023), quando auxiliam na coleta de dados, no processamento de informações, a partir de imagens aéreas (drones ou satélites). Seu uso mais comum é na coleta direta de dados para gerar diagnósticos e decidir sobre a aplicação de insumos, como fertilizantes, reguladores de água e solo, todos em taxas diferentes, definindo a dosagem ideal. Com o surgimento de novos dispositivos de IoT em soluções de AP, novas opções de sensores e aplicações estão sendo ofertadas (Tabela 8). Segundo Shukla et al. (2023), as perspectivas futuras do uso de sensores na agricultura de precisão apontam para processamento distribuído em soluções embarcadas, disponibilizando informações mais precisas por modelagem preditivas por inteligência artificial, com transferências de dados muito mais rápidas, em baixa latência, pela adoção da tecnologia 5G.

Tabela 8: Sensores e suas funcionalidades na solução de AP. Fonte: Adaptado de Singh et al. (2020).

Sensor	Tecnologia	Funcionalidades	Integração
Sensor de Umidade do Solo	IoT e Sensoriamento Capacitivo	Monitora o nível de umidade do solo em tempo real	Sistemas de irrigação automatizada via IoT
Câmera Multiespectral	Sensoriamento Remoto Passivo	Captura imagens multiespectrais para cálculo de índices vegetativos (NDVI, EVI) e avaliação da saúde da vegetação	Integração com drones, satélites e IA para análise de culturas
Câmera Hiperespectral	Sensoriamento Remoto Passivo/Ativo	Detecta variações espectrais detalhadas da vegetação, solo e água	Integração com drones, IA e GIS para mapeamento e análise preditiva
Câmera RGB	Sensoriamento Remoto Passivo	Captura imagens visíveis para detecção de pragas e monitoramento da colheita	Integração com drones e IA para análise de imagens agrícolas
Sensor de Condutividade	Eletromagnética	Avalia a variabilidade de nutrientes e salinidade do solo	Integração com sistemas de VRT para otimização de fertilizantes

Elétrica do Solo

Lidar	Sensoriamento Remoto Ativo (Laser)	Mapeia a topografia do terreno e altura das plantas	Integração com drones e GIS para geração de modelos 3D
Estação Meteorológica	IoT e Sensores Climáticos	Monitora clima, temperatura, umidade, vento e precipitação	Integração com plataformas de previsão meteorológica e controle de irrigação
Sensores de Fertilizante	IoT e NIR	Monitora os níveis de nutrientes no solo e aplica fertilizantes de acordo	Integração com tratores e pulverizadores de taxa variável
Sensor de Nível de Água	Ultrassom e IoT	Monitora os níveis de água em reservatórios e sistemas de irrigação	Integração com sistemas de controle de irrigação automatizada
Sensor de PH do Solo	Eletroquímico e IoT	Mede o nível de acidez do solo em tempo real	Integração com software de gestão de fertilidade do solo
Radar de Abertura Sintética (SAR)	Sensoriamento Remoto Ativo (Micro-ondas)	Captura dados em todas as condições climáticas para monitoramento do solo e culturas	Integração com satélites e plataformas GIS para análise geoespacial

Veículos automatizados autômatos

A produção industrial avança para integração de hardware com alta tecnologia, inclusive nos equipamentos e dispositivos eletrônicos utilizados no agronegócio, como tratores, colheitadeiras, semeadoras e irrigadores. As soluções atuais de AP disponibilizadas pelas agtechs já suportam integração de veículos agrícolas automatizados com os demais componentes (Raghunath e Devi, 2022; Mukati et al., 2024), operados por softwares aplicativos embarcados e outros dispositivos de controle (Chakraborty et al., 2022; Sivasangari et al., 2023), auxiliados em navegação no campo por GPS, redes sem fio, RFID (Vrchota et al., 2022; Ünal et al.,

2023), integrando-se a dispositivos de IoT e georreferenciamento por drones ou satélites (Padmanabhan e Megalingam, 2022), para coletar dados em tempo real, minimizando intervenção humana local.

Dentre as vantagens de sua utilização para atividades no campo (preparação solo, plantio, irrigação e colheita e transporte), estão a redução de mão de obra de campo, eficiência e produtividade, através da rapidez e precisão na intervenção segmentada de ocorrências de doenças, pragas, deficiências de solo e plantas, estresse hídrico e colheita e transporte, frente aos diagnósticos e ajustes necessários identificados (Kumari e Bhavanishankar, 2023). Entre as

desvantagens e problemas identificados, dada a localização da maioria dos plantios em áreas rurais, estão os custos para aquisição por importação ou manutenção altos, necessidade de pessoal especializado, gerando barreira de utilização para médios e grandes produtores rurais (Shukla et al.,

2023). Dentre os desafios encontrados por sua utilização em soluções de AP (Tabela 9), estão a complexidade de integração dos diversos componentes envolvidos, além da dificuldade de financiamentos (Sivasangari et al., 2023; Emmi et al., 2023).

Tabela 9: Principais veículos agrícolas utilizados na solução de AP. Fonte: Autores (2023).

Tipo de Veículo	Finalidade/uso	Componentes de AP Associados	Fabricantes	Custos estimados (Brasil)	Desafios atuais
Tratores Automatizados	Plantio, colheita, pulverização (diversos índices de plantio, solo)	Sensores solo, GPS, RFID, IoT, IA, big data	John Deere, Case IH	US\$ 600k	Necessidade de cobertura de rede eficiente. Com alto custo de manutenção
Colheitadeiras	Colheita autônoma de grãos com reconhecimento inteligente	SR Lidar, Sensores solo, VRT, wifi, IoT, IA, GPS	CLAAS, New Holland	US\$ 300k - \$500k	Alto custo de aquisição
Semeadoras	Semeadura de insumos diversos com precisão	Sensores de plantio, VRT, IoT, IA, GPS	John Deere, Kinze	US\$ 100k - \$200k	Complexidade na calibração dos sensores
Irrigadores Automatizados	Distribuição de água com dosagem inteligente	Sensores de umidade, VRT, IA, IoT, GPS	Valmont, Lindsay Corp	US\$ 50k - \$100k	Manutenção regular, falhas de sensores
Capinadores Robóticos	Identificação, remoção de ervas daninhas diversas	Visão computacional, IA, Sensores de solo, GPS	Naïo Technologies	US\$ 50k - \$150k	Dificuldade em condições adversas do solo

Tecnologia de blockchain

A proposta para utilização da tecnologia de blockchain (Gangwani et al., 2023) nas soluções de agricultura de precisão pelas agtechs reflete as funcionalidades já implementadas na indústria (Javaid et al., 2021; Ali-Khusein e Urquhart, 2023). A finalidade principal de seu uso está ligada de suportar aplicações P2P com transações descentralizadas quanto a armazenamento, compartilhamento e tratamento de dados agrícolas, com alta proteção e confiabilidade (Navdeep et al., 2022), facilitando também a implementação de outras tecnologias, como computação em nuvem, virtualização de dados e processamento

distribuído, em aplicativos móveis. A sua implementação propiciará a rastreabilidade de forma segura e transparente de dados, facilitando as propostas de financiamento de plantios por diversas entidades do mercado (Figura 9). A tecnologia de blockchain pode também contribuir para minorar desperdícios de produtos na cadeia logística. As agtechs e empresas de pesquisas estudam possibilidade de implementação da tecnologia em dispositivos embarcados, como nas soluções de IoT e sensores (Padhy et al., 2023), facilitando novas funcionalidades na solução, como contratos inteligentes de compra e venda, precificação e transações de pagamentos de fornecedores e compradores (Pakseresht et al.,

2023). Mas com certeza, a incorporação da tecnologia de blockchain irá aumentar a confiabilidade das aplicações de AP na geração dos dados da produção agrícola, como qualidade dos produtos, previsões de resultados da produção e segurança alimentar, pelo controle autenticado de insumos e agrotóxicos, além de aumentar segurança para análises com privacidade de dados

para sustentabilidade (Adewusi et al., 2024) do setor. Em um artigo de revisão, Pakseresht et al. (2023) analisam 44 artigos, e resumem as principais atividades relacionadas com agricultura empregam a tecnologia blockchain (Figura 8), demonstrando a tendência de desenvolvimento de sistemas voltados para propiciar o rastreamento da produção agrícola.

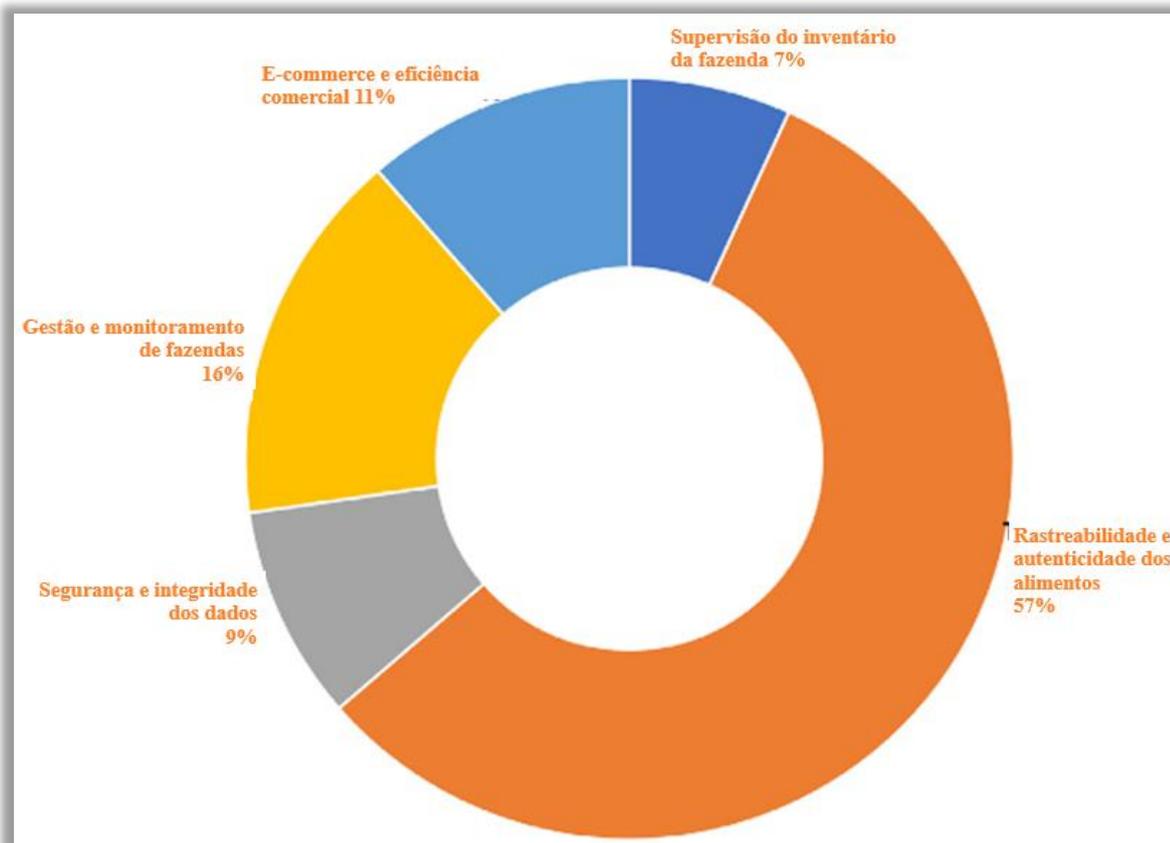


Figura 9: Distribuição de casos de uso de blockchain em agroalimentação. Fonte: Adaptado de Pakseresht et. (2023).

Mobile Apps

As aplicações móveis em soluções de AP foram inicialmente introduzidas para disponibilizar versões customizadas da AP, principalmente para o gerenciamento em campo. Dentre as funcionalidades básicas existentes: Interação com sensores, navegação de drones, redes wifi, informações de SR e IoT, diagnósticos por seções do terreno, disponibilizando informações de gerenciamento agrícola; recurso de informações agrícolas; calculadora agrícola; notícias agrícolas; clima; e m-governo (Colucci et al., 2022; Nagy e

Lakatos, 2022). Os dispositivos móveis atuais possuem alta capacidade de processamento, permitindo combinar suporte a transações locais, oriundas de sensores do aparelho celular, facilitando utilização de computação distribuída, com integração dos dados coletados com componentes como big data, realizando armazenamento em nuvem (computação em nuvem) pela aplicação de AP. Além disso, propiciam aumento da conectividade dos diversos componentes da solução, por oferta da cobertura da

rede celular. Dentre as principais operações desenvolvidas pelas agtechs, estão (Tabela 10): Cultivo; gerenciamento de fazenda; e sistema de informação. Novas características estão sendo

incorporadas, facilitadas por tecnologias como computação em nuvem, computação distribuída, computação de borda e redes celulares 5G (Tomaszewski et al., 2022).

Tabela 10: Aplicações móveis na solução de AP. Fonte: Adaptado de Mendes et al. (2020).

Módulo da solução de AP	Finalidade	Aplicações suportadas
Operações de Cultivo	Proteção e Diagnóstico	* Detecção e diagnóstico de pragas e doenças; * Identificação e tratamento de ervas daninhas; * Diagnóstico do solo e das plantas
	Nutrição e Fertilização	* Monitoramento da nutrição das culturas; * Gestão da pulverização; * Aplicação de fertilizantes.
	Irrigação	* Monitoramento do status hídrico das culturas; * Suporte à decisão de irrigação.
	Crescimento e Gestão do Dossel Vegetal	* Acompanhamento do crescimento do dossel; * Cálculo do Índice de Área Foliar (IAF).
	Colheita	* Estimativa de produtividade; * Indicadores de qualidade.
Gestão da Fazenda	Mapeamento de Campo e Informações do Solo	* Cálculo de localização e área do campo; * Identificação de pontos de coleta de amostras; * Indicadores agrícolas do solo (cor, pH, NPK, conteúdo de carbono).
	Gestão de Máquinas	* Estimativa de custos de máquinas; * Monitoramento em tempo real de trajetórias no campo; * Monitoramento de atividades, produtividade e uso eficiente das máquinas.
	Controle de Atividades da Fazenda	* Gestão de tarefas no campo * Gerenciamento de atividades dos trabalhadores rurais
Sistema de Informação	Informações e Insights	* Dicas e conhecimentos agrícolas; * Informações de mercado; * Notícias relevantes; * Comunicação com especialistas; * Dados climáticos.

Resultados e discussão

Os recursos investidos no mercado global para soluções de agricultura de precisão foram estimados em US\$ 5,49 bilhões em 2021, projetado

para superar US\$ 19 bilhões até 2030, com CAGR (taxa de crescimento anual composta) de 14,95% de 2022 (Nair,2023), onde analistas do MarketsandMarkets (2024) projetaram um

crescimento de US\$ 81,5 bilhões em 2023, entre os dez maiores mercados produtores agrícolas, indo a US\$ 94,23 bilhões em 2024. Estes números demonstram que a solução de AP tem ganhado reconhecimento, ampliando sua base instalada, resultado obtido de sua utilização nos mais variados tipos de plantio, em diversas regiões, contribuindo para a produção agrícola mundial

(Figura 10). Embora diversos trabalhos de pesquisas apontem que as soluções de AP auxiliem na produção agrícola, ainda há poucos relatórios abrangentes publicados, detalhando efetivamente a contribuição para a produção mundial, por região/país e cultura, frente a outros métodos de promoção desse crescimento.

País/Área	Evolução (produção em toneladas)		Mudança Absoluta	Mudança Relativa
	2009	2022		
Brasil	50.719.824	109.420.720	↑ 58.700.896	↑ 116%
Índia	16.719.500	33.729.540	↑ 17.010.040	↑ 102%
China (FAO)	164.107.550	277.415.550	↑ 113.308.000	↑ 69%
EUA	331.921.060	348.750.940	↑ 16.829.880	↑ 5%
América do Sul (FAO)	74.608.780	183.039.840	↑ 108.431.060	↑ 145%
Ásia (FAO)	234.263.360	389.892.540	↑ 155.629.180	↑ 66%
África (FAO)	60.750.668	92.800.970	↑ 32.050.302	↑ 53%
América Central (FAO)	23.830.832	30.717.768	↑ 6.886.936	↑ 29%
Europa (FAO)	84.589.620	102.694.190	↑ 18.104.570	↑ 21%
América do Norte	365.964.160	394.432.030	↑ 28.467.870	↑ 8%
Oceania (FAO)	627.736	637.819	↑ 10.083	↑ 2%

Figura 10: Evolução da produção mundial de milho por países e regiões, de 2009 a 2022. Fonte: Adaptado de FAO-STATS (2022).

No mesmo período analisado para a produção mundial de milho (Figura 10), dados históricos da evolução do uso de terra por países produtores mundiais (Figura 11), demonstram que países como China, Brasil e EUA apresentaram crescimento significativo de sua produção agrícola, contrastando proporcionalmente com o aumento da área plantada, sugerindo uso mais eficiente da produtividade por hectare plantado. Coincidentemente, estes países estão entre os que

mais produtores têm investido na utilização das soluções de AP, embora não se possa corroborar essa constatação com dados disponíveis em pesquisas na literatura. Dos relatórios citados acima, é possível evidenciar, corroborado com o relatório da FAO (FAO-STATS, 2022), os países mais desenvolvidos apresentam diferencial competitivo pelo reconhecimento da qualidade de produtos maior produtividade por hectare plantado.

País/Área	Evolução (uso terras em hectares)		Mudança	
	2009	2022	Absoluta	Relativa
Brasil	13.654.715	21.037.668	↑ 7.382.953	↑ 54%
China (FAO)	31.203.368	43.102.264	↑ 11.898.896	↑ 38%
Índia	8.261.600	9.957.950	↑ 1.696.350	↑ 21%
EUA	32.168.810	32.054.280	↓ -114.530	↓ 0%
América do Sul (FAO)	19.606.748	33.325.184	↑ 13.718.436	↑ 70%
África (FAO)	30.301.654	41.770.892	↑ 11.469.238	↑ 38%
Ásia (FAO)	53.481.412	68.207.820	↑ 14.726.408	↑ 28%
Europa (FAO)	13.932.788	17.533.920	↑ 3.601.132	↑ 26%
América Central (FAO)	8.058.159	8.684.168	↑ 626.009	↑ 8%
América do Norte	41.979.704	42.558.610	↓ 578.906	↓ 1%
Oceania (FAO)	90.728	73.583	↑ -17.145	↑ -19%

Figura 11: Evolução do uso de terras para a produção mundial de milho, de 2009 a 2022. Fonte: FAO-STATS (2022).

Este artigo contextualizou a importância das aplicações de agricultura de precisão para a produção mundial de alimentos (Memon et al., 2023; Barbosa-Júnior et al., 2024; Bazargani e Deemyad, 2024; Govindaraj et al., 2024; Maurya et al., 2024; Petrović et al., 2024), ressaltando a importância de sua utilização por produtores nos países desenvolvidos, e percentualmente maior entre os grandes produtores agrícolas no mercado de países em desenvolvimento (Nair et al., 2023), impactando na melhoria da gestão do negócio (Szira et al., 2023), aumento da produtividade (Rimpika et al., 2023), otimização do monitoramento (Nagy e Lakatos, 2022) e gerenciamento das variáveis envolvidas do ciclo de crescimento e colheita da cultura (Upadhyay et al., 2024).

Outro aspecto que o trabalho procurou enfatizar focou na descrição funcional dos principais componentes constituintes da solução (Alahe et al., 2024; Canicatti e Vallone, 2024; Ijaz e Wozniak, 2024; Kumar, Sharma, Kedam, Patel, & Rathnayake, 2024a; Lamb et al., 2022; Lanucara et al., 2024; Muhammed et al., 2024; Padhiary et al., 2024; Raihan, 2024; Sharma e Shivandu, 2024; Vinod et al., 2024; Wang et al., 2024;), atualizando sobre a importância destes elementos, e a necessidade de sua perfeita integração na solução. A revisão apontou que a maioria das publicações selecionadas, está ligada à descrição geral da solução, descrição funcional de seus componentes,

principalmente os mais antigos, que evoluíram com as soluções de AP, como mapas multiespectrais de índices de solo, plantas e recursos hídricos, gerados por sensoriamento remoto, controle e aplicação de insumos por VRT, sensores de campo para métricas, aplicativos de SIG, e sistema de GPS para geolocalização e aeronavegação. Mais recentemente, cresceram publicações com dispositivos de IoT, sistema de informações com *big data*, drones, WebSIG, sensores inteligentes, sistemas automatizados, aplicativos móveis, redes 5 G, métodos de tratamento de dados e geração de informações estratégicas, com aprendizado de máquina (IA). Agora, intensificam-se artigos com preocupação em segurança e privacidade das transações do negócio, através de novos protocolos de comunicação e criptografia (Ongadi, 2024). Assuntos como regulamentações, políticas públicas de financiamento e fomento, atores envolvidos são dispersos, com pouca profundidade ou desatualizados.

Os benefícios citados na maioria dos artigos pesquisados nesta pesquisa focam na produtividade agrícola, gestão do negócio, principalmente relacionados com o controle de custos (preparação solo, aplicação de insumos e agroquímicos, uso de água, e energia); monitoramento (predição de colheita, logística e precificação); segurança e qualidade alimentar dos produtos gerados (detecção doenças e pragas, avaliação de nutrientes de solo e plantas, dosagem

precisa de agrotóxicos), além da sustentabilidade do negócio (época de plantio, roteiro de plantio, rastreabilidade). Outro fato destacado nos artigos é o constante aperfeiçoamento dos componentes da solução de AP, oriundo da diversificação da oferta de sistemas aplicativos por novas agtechs, que propiciará gradual redução de custos, combinada com a evolução de tecnologias decorrentes da indústria, e da ciência da computação, que viabilizarão a implementação da AP em larga escala, com maior modularidade, permitindo otimização das predições e monitoramento, suportados por ferramentas de inteligência artificial, convergindo para o conceito da agricultura 5.0 (Ragazou et al., 2022; Wrzecińska et al., 2023). O aperfeiçoamento das soluções de AP, mediante redução dos problemas de integração entre componentes, também trará benefícios para os produtores usuários, pois aumentará a modularidade, possibilitando maior personalização ao tipo de cultura e tamanho da produção, concentrando a automação nas fases mais importantes do processo produtivo em uma fazenda. Por fim, o ganho percebido decorrente do aperfeiçoamento dos sistemas aplicativos de AP, com incorporação de requisitos para controlar com mais rigor os danos ambientais, resultante da crescente exigência do conceito de sustentabilidade ambiental (Azlan et al., 2024), pressionando por mais pesquisas em direção de temas como busca substitutivos de agroquímicos (biopesticidas), maior eficiência no uso de agrotóxicos, além do uso de energia limpa, agricultura sustentável, segurança alimentar e rastreabilidade de produtos.

Questões versando sobre problemas e dificuldades relacionadas à solução de AP, inerentes à barreira de ampliação da solução nos diversos mercados mundiais, também foram abordadas. Pesquisas sobre publicações na literatura analisadas, ressaltaram problemas ligados, principalmente, na integração de tecnologias utilizadas na solução de AP, complexidade operacional percebida por fazendeiros, falta de suporte especializado local. Além disso, a falta de soluções personalizadas, aderentes ao seu cenário de produção local, denota que o foco das agtechs (Mendes et al., 2022) e políticas em geral, nos países em desenvolvimento, ainda é voltado para grandes clientes, sem, entretanto, estimular a disponibilização de soluções modulares, gerando indecisão entre pequenos e médios fazendeiros (John et al., 2023), sobre a real possibilidade de retorno sobre investimentos, dado os preços altos de aquisição e manutenção da

solução. Outra questão ressaltada foi a de dificuldades de financiamento, tanto para aquisição da solução e equipamentos, quanto para capacitação de pessoal, quanto para serviços de suporte local especializado, visando sua manutenção, dependendo da região/país.

Como pontos de atenção sobre os desafios atuais para ampla aceitação da agricultura de precisão, duas vertentes foram citadas nesta revisão:

(a) Uma sobre necessidade de melhoria de planos de ações a serem procedidos, visando conscientização de governos e entidades ligadas a seu desenvolvimento, notadamente por nações em desenvolvimento, de forma que estimulem a ampliação de investimento públicos e privados, para amadurecimento e fortalecimento de base tecnológica de pesquisas e desenvolvimento e de serviços, atentando para quão estratégico será intensificar condições de apoio ao crescimento de segmento local especializado, representados por agtechs, programas de capacitação de serviços de IT, linhas de pesquisas em universidades e centros de tecnologia, incremento na diversificação de entidades de apoio aos pequenos e médios produtores, visando minimizar aceitação de seu uso, impactando diretamente no PIB, geração de empregos, e na questão ambiental. China, Índia e países da Europa Oriental são exemplos para esta questão, investindo pesadamente para superar estes desafios (Rimpika et al., 2023);

(b) Outra vertente é observada pela lacuna da competição por mercados consumidores, existente entre grandes países produtores desenvolvidos e em desenvolvimento, quanto à exportação de seus produtos (FAO-STATS, 2022; Nair, 2023), afetando lucratividade, causada nos primeiros, pela rapidez de incorporação das inovações tecnológicas nas soluções de AP (figura 2), promovidas pela variedade de agtechs altamente especializadas, pioneiras no lançamento de novas funcionalidades, advindas da disponibilidade de forte base tecnológica instalada, com políticas agressivas de fomento, minimizando custos, que diferenciam as soluções de AP, afetando a qualidade da gestão do negócio, otimização de recursos e custos, qualidade nutricional e segurança alimentar relacionadas a estes produtos, com maior precisão de projeções e comercialização de produtos. A diminuição desta lacuna para países produtores em desenvolvimento, passa por avaliar e adotar medidas, se for o caso, como nos exemplos da China, Índia e países da Europa Oriental, buscando diferenciação nos

gradativamente sendo incorporadas nas soluções de AP, e permitirão implementar futuros mecanismos de auditoria, inspeções e fiscalizações mais efetivas, melhorando a percepção de qualidade nutricional e segurança alimentar dos produtos gerados, abrindo caminho para

alternativas, como o uso de energia renovável (Rehman et al., 2024), bioinsumos (Biotec, 2024), e selo verde, visando redução de danos ambientais nos maiores produtores mundiais (Kanwal et al., 2022; Zanin et al., 2022).

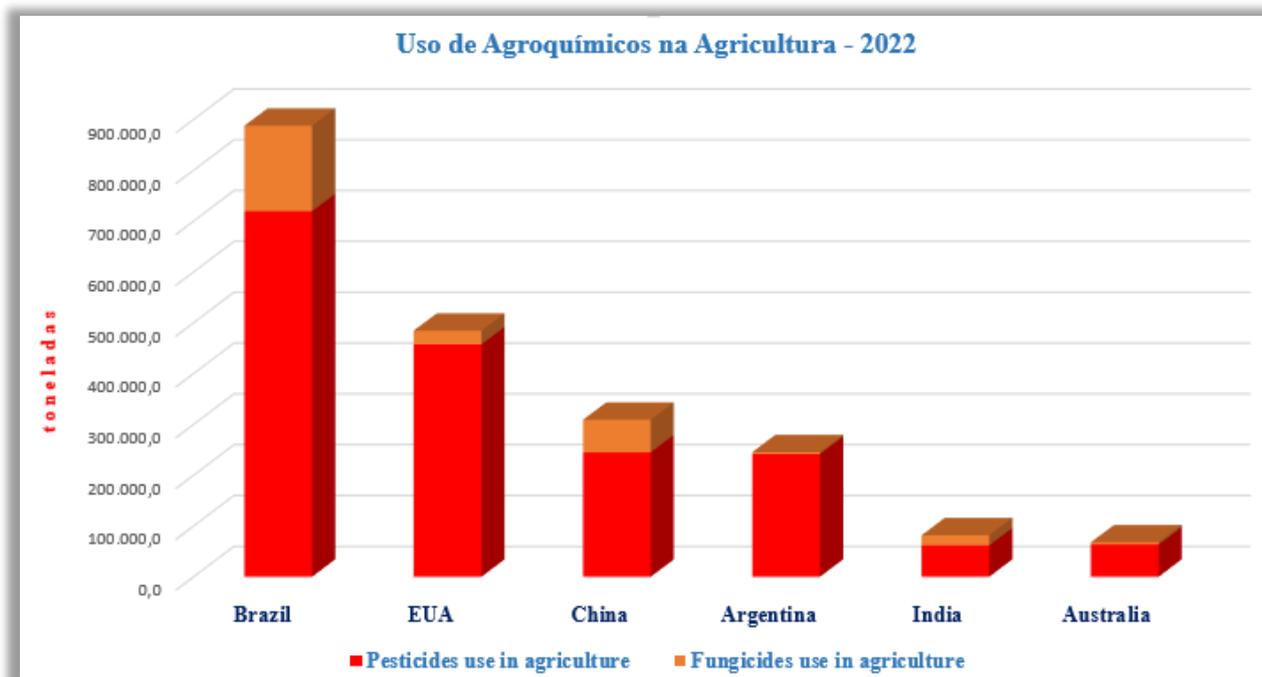


Figura 12: Emprego de agrotóxicos por alguns dos maiores produtores mundiais. Fonte: Adaptado de FAO-STATS (2022).

As perspectivas futuras para soluções de AP são promissoras e apontam, após pesquisas na literatura, para alguns fatores-chave, como:

(a) Agtechs disponibilizarão novas versões de sistemas aplicativos modulares aos projetos de AP, possibilitando a abstração de sua complexidade e minimização de barreira de entrada, principalmente para pequenos e médios produtores, que não contam com suporte local de pessoal especializado para operar e manter estes sistemas;

(b) Tendências de redução de custos iniciais e manutenção, pela oferta de sistemas mais personalizados e flexíveis pelas agtechs, além da ampliação da oferta de crédito e fomento, em termos mundiais, principalmente em países como China, Índia, Brasil e da Europa Central;

(c) Aumento da capacidade de processamento de grandes volumes de dados mediante novas tecnologias, como: Computação em nuvem (*cloud computing*), virtualização de sistemas (*system virtualization*), técnicas de

inteligência artificial com aprendizado de máquina profundo (*deep learning*, Sanaeifar et al., 2023), surgimento de novos protocolos de comunicação, como LTE-M, 6G (Polymeni et al., 2023), e NB-IoT (Soussi et al., 2024), bancos de dados interativos (*datalakes*), que suportarão implementação de novos dispositivos baseados em computação embarcada e distribuída (*edge computing*), como sensores, drones, veículos autômatos de campo, e aparelhos móveis para navegação amigável (*mobile handsets*);

(d) Melhor acurácia na predição de análises e predições do agronegócio, incrementando novas variáveis ao monitoramento e controle do plantio e colheita, pelo melhoramento de insumos com biotecnologia (Gor et al., 2024; Daware et al., 2024), nanotecnologia (Zain et al., 2024; Miguel-Rojas e Pérez-de-Luque, 2024) robotização inteligente (Cheng et al., 2023), que tornarão mais eficientes os procedimentos de identificação de doenças, deficiências de plantas e solo, dosagem de

insumos e agrotóxicos (Kumar, Faoq, & Qureshi, 2024b).

Conclusão

Sistemas aplicativos de agricultura de precisão (AP) nos diversos temas tratados nesta revisão mostraram que a ferramenta está em evolução, com potencial significativo para contribuir para o aumento da produção mundial de alimentos, ao tempo que melhora a mitigação de seus impactos ambientais negativos.

Com a análise detalhada dos principais componentes de uma solução típica de agricultura de precisão, baseada em pesquisas na literatura de textos de revisão, os diversos problemas, desafios e benefícios, demonstram o potencial para aumentar sua taxa de penetração, principalmente em países em desenvolvimento, foram discutidos. O direcionamento das principais tendências requeridas para solução, também foi esclarecido, mostrando lacunas ainda não preenchidas, que podem ser atendidas por sua evolução tecnológica, pela disponibilidade de infraestrutura e por ações governamentais e privadas, visando diminuir a barreira de altos custos de implementação, principalmente para pequenos e médios produtores.

A revisão também ressaltou que, mesmo com maior utilização pelos países produtores mais desenvolvidos, além dos vultuosos valores financeiros investidos no desenvolvimento de soluções de AP, sua taxa de penetração média é baixa, em termos mundiais, havendo necessidade de readequação ou aperfeiçoamento das políticas de financiamento existentes para desenvolvedores (agtechs) e clientes, com aumento de ações de fomento de polos de capacitação e formação de profissionais e pesquisadores locais.

Este trabalho também aponta para a necessidade de mais pesquisas voltadas à personalização das soluções de AP, de forma a torná-las mais acessíveis e funcionais, em diferentes contextos geográficos e econômicos. Isto pode ser atingido com a inovação em tecnologias como IA, IoT e biotecnologia, que prometem transformar o setor agrícola, contribuindo não apenas para aumento da produtividade por hectare plantado, retardando a expansão de áreas plantadas sob diversos biomas. Também podem auxiliar para a segurança alimentar global, culminando na mitigação de seus efeitos para as mudanças climáticas. Em última análise, a solução AP pode desempenhar um papel crucial na agricultura sustentável, com sua evolução para o conceito de agricultura digital 5.0, devendo ser considerada prioridade estratégica em políticas agrícolas ao redor do mundo, combinada, ou alternativamente a outras medidas de contenção de danos ambientais.

Espera-se, com os temas aqui discutidos, contribuir para a pesquisa futura, que pode se concentrar em integrar essas tecnologias emergentes, a problemática social em torno da implantação deste tipo de solução, a criação de mecanismos para avaliação de seus impactos na economia, no meio ambiente, de forma a capacitar países produtores agrícolas, principalmente os em desenvolvimento, crescimento de sua riqueza, maior disponibilidade de alimentos, com processos sustentáveis.

Referências

- Aashu, Rajwar, K., Pant, M., & Deep, K. (2024). Application of Machine Learning in Agriculture: Recent Trends and Future Research Avenues (No. arXiv:2405.17465). arXiv. <http://arxiv.org/abs/2405.17465>.
- Abdelbaki, A., & Udelhoven, T. (2022). A Review of Hybrid Approaches for Quantitative Assessment of Crop Traits Using Optical Remote Sensing: Research Trends and Future Directions. *Remote Sensing*, 14(15), Artigo 15. <https://doi.org/10.3390/rs14153515>.
- Abobatta, W. F. (2021). Precision Agriculture: A New Tool for Development. Em S. M. Abd El-Kader & B. M. Mohammad El-Basioni (Orgs.), *Advances in Environmental Engineering and Green Technologies* (p. 23–45). IGI - Global. <https://doi.org/10.4018/978-1-7998-5000-7.ch002>

- Abu, N. S., Bukhari, W. M., Ong, C. H., Kassim, A. M., Izzuddin, T. A., Sukhaimie, M. N., Norasikin, M. A., & Rasid, A. F. A. (2022). Internet of Things Applications in Precision Agriculture: A Review. *Journal of Robotics and Control (JRC)*, 3(3), Artigo 3. <https://doi.org/10.18196/jrc.v3i3.14159>
- Adewusi, Onyeka Franca Asuzu, Temidayo Olorunsogo, Temidayo Olorunsogo, Ejuma Adaga, & Donald Obinna Daraojimba. (2024). AI in precision agriculture: A review of technologies for sustainable farming practices. *World Journal of Advanced Research and Reviews*, 1, 2276–2285. <https://doi.org/10.30574/wjarr.2024.21.1.0314>
- AgFunder. (2022). 10 Years of AgriFoodTech (and AgFunder!) | A Celebration and Potted History. <https://AgFundernews.com/>. Disponível em: <https://agfunder.com/research/10-years-agrifoodtech-and-agfunder-celebration-and-potted-history/>. Acesso em: 13 jan. 2022.
- Ahmed, A., Hamza, S., Tahir, M. N., Saleem, S. R., & Ahmed, R. (2022). A Performance Comparison of Variable Rate Technologies for Spot-Specific and Uniform Spraying for Citrus Orchard. *The 1st International Precision Agriculture Pakistan Conference 2022 (PAPC 2022) – Change the Culture of Agriculture*, 21. <https://doi.org/10.3390/environsciproc202203021>
- Akhtar, M. N., Shaikh, A. J., Khan, A., Awais, H., Bakar, E. A., & Othman, A. R. (2021). Smart Sensing with Edge Computing in Precision Agriculture for Soil Assessment and Heavy Metal Monitoring: A Review. *Agriculture*, 11(6), Artigo 6. <https://doi.org/10.3390/agriculture11060475>
- Alahe, M. A., Wi, L., Chang, Y., Gummi, S. R., Kermeshi, J., & Yang, K. (2024). Cybersecurity in smart agriculture: Threat types, current status and future trends. *Computers and Electronics in Agriculture Journal*, 226. <https://doi.org/10.1016/j.compag.2024.109401>
- Alahmad, T., Neményi, M., & Nyéki, A. (2023). Applying IoT Sensors and Big Data to Improve Precision Crop Production: A Review. *Agronomy*, 13(10), Artigo 10. <https://doi.org/10.3390/agronomy13102603>
- Arya, N., Sharma, A., & Vajpayee, A. (2024). Crop Disease Identification Using Deep Learning Techniques. *International Journal of Innovative Science and Research Technology (IJISRT)*, 946–952. <https://doi.org/10.38124/ijisrt/IJISRT24MAR885>
- Azlan, Z. H. Z., Junaini, S. N., Bolhassan, N. A., Whai, R., & Arip, M. A. (2024). Harvesting a sustainable future: An overview of smart agriculture's role in social, economic and environmental sustainability. *Journal of Cleaner Production*, 434. <https://doi.org/10.1016/j.jclepro.2023.140338>
- Barbosa-Júnior, M. R., Rodrigues, B. R. de A., Carreira, V. dos S., Brito Filho, A. L., & Shiratsuchi, L. S. (2024). Precision agriculture in the United States: A comprehensive meta-review inspiring further research innovation and adoption. *Computers and Electronics in Agriculture*, 221. <https://doi.org/10.1016/j.compag.2024.108993>
- Bazargani, K., & Deemyad, T. (2024). Automation's impact on agriculture: Opportunities, challenges and economic effects. *Robotics*, 13(2), 33. <https://doi.org/10.3390/robotics13020033>
- Belcore, E., Angeli, S., Colucci, E., Musci, M. A., & Aicardi, I. (2021). Precision agriculture workflow from data collection to data management using FOSS tools: An application in northern Italy vineyard. *International Journal of Geo-Information*, 10. <https://doi.org/10.3390/ijgi10040236>
- Bhat, S. A., & Huang, N.-F. (2021). Big data and AI revolution in precision agriculture: Survey and challenges. *IEEE Access*, 9, 110209–110222. <https://doi.org/10.1109/ACCESS.2021.3102227>
- Biotech. (2020). The National Center for Genetic Engineering and Biotechnology (Biotech). Disponível em: <https://www.biotech.or.th/home/en/overview-en/>. Acesso em: 13 ago. 2024.

- BNDES - BANCO NACIONAL DO DESENVOLVIMENTO. (2022). Programas de financiamento do Governo Federal do Brasil. Disponível em: <https://www.bnb.gov.br/>. Acesso em: 21 jun. 2022.
- Bodrud-Doza, M., Yang, W., Miranda, R. de Q., Martins, A., DeVries, B., & Fraser, E. D. G. (2023). Towards implementing precision conservation practices in agricultural watersheds: A review of the use and prospects of spatial decision support systems and tools. *Science of the Total Environment*, 905. <https://doi.org/10.1016/j.scitotenv.2023.167118>
- Bomoi, M. I., Nawari, N. M., Abd Aziz, S., & Mohd Kassim, M. S. (2022). Sensing technologies for measuring grain loss during harvest in paddy field: A review. *AgriEngineering*, 4(1), 292–310. <https://doi.org/10.3390/agriengineering4010020>
- Boopathi, S. (2024). Sustainable development using IoT and AI techniques for water utilization in agriculture. *IGI Global*, <https://doi.org/10.4018/979-8-3693-1722-8.ch012>
- Bramantoro, A., Suhaili, W. S., & Siau, N. Z. (2022). Precision agriculture through weather forecasting. 2022 International Conference on Digital Transformation and Intelligence (ICIDI), IEEE Journal. <https://doi.org/10.1109/icdi57181.2022.10007299>
- Brini, M. (2023). AgTech adoption by farm's segment and production phase. ResearchGate. https://www.researchgate.net/publication/370004655_AgTech_adoption_by_farm's_segment_and_production_phase_Digital_agriculture_and_farming_phases#fullTextFileContent
- Bwambale, E., Abagale, F. K., & Anornu, G. K. (2022). Smart irrigation monitoring and control strategies for improving water use efficiency in precision agriculture: A review. *Agricultural Water Management*, 260, 107324. <https://doi.org/10.1016/j.agwat.2021.107324>
- Canicatti, M., & Vallone, M. (2024). Drones in vegetable crops: A systematic literature review. *Agricultural Technology*, 7. <https://doi.org/10.1016/j.atech.2024.100396>
- CEBDS - Conselho Empresarial Brasileiro para o Desenvolvimento Sustentável. (2021). Sistemas agroalimentares. CEBDS. Disponível em: <https://cebds.org/camara-tematica/sistemas-agroalimentares/>. Acesso em: 14 jun. 2022
- Chaim, A., Neto, J. C., Queirós, L. R., & Junior, A. L. (2014). AgroTIC em agricultura de precisão e automação agrícola, capítulo 10. Disponível em: <https://www.alice.cnptia.embrapa.br/alice/handle/doc/1040539>. Acesso em: 21 jul. 2022
- Chakraborty, S., Elangovan, D., Govindarajan, P. L., ELnaggar, M. F., Alrashed, M. M., & Kamel, S. (2022). A comprehensive review of path planning for agricultural ground robots. *Sustainability*, 14(15), 9156. <https://doi.org/10.3390/su14159156>
- Cheng, C., Fu, J., Su, H., & Ren, L. (2023). Recent Advancements in Agriculture Robots: Benefits and Challenges. *Machines*, 11(1), 48. <https://doi.org/10.3390/machines11010048>
- Chintakunta, A. N., Koganti, S., Nuthakki, Y., & Kolluru, S. (2023). Deep learning and sustainability in agriculture: A systematic review. *International Journal of Computer Science & Mobile Computing*, 12(8), 150-164. <https://doi.org/10.47760/ijesmc.2023.v12i08.017>
- Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2012). SciMAT: A new science mapping analysis software tool. *Journal of the American Society for Information Science and Technology*, 63(8), 1609-1630. <https://doi.org/10.1002/asi.22688>
- Colucci, G., Botta, A., Tagliavini, L., Cavallone, P., Baglieri, L., & Quaglia, G. (2022). Kinematic modeling and motion planning of the mobile manipulator Agri.Q for precision agriculture. *Machines*, 10(5), 321. <https://doi.org/10.3390/machines10050321>
- Das, P., Kumar, A., Nath, N., Mishra, L., & Gautam, V. (2023). Application of drone technology in precision agriculture: Enhancing efficiency and sustainability. Em *Book, Advances in Land and Water Resources Management Concept and Application* (p. 113-124). <https://www.researchgate.net/publication/373>

833447_application_of_drone_technology_in_precision_agriculture_enhancing_efficiency_and_sustainability

Daware, V., Mondal, K., & Khandagale, S. (2024). Precision agriculture and biotechnology. ResearchGate. https://www.researchgate.net/publication/378343059_precision_agriculture_and_biotechnology

Debauche, O., Mahmoudi, S., Manneback, P., & Lebeau, R. (2022). Cloud and distributed architectures for data management in agriculture 4.0: Review and future trends. *Journal of King Saud University – Computer and Information Sciences*, 1319-1578. <https://doi.org/10.1016/j.jksuci.2021.09.015>

Dos Santos, R. P., Fachada, N., Beko, M., & Leithardt, V. R. Q. (2023). A rapid review on the use of free and open-source technologies and software applied to precision agriculture practices. *Journal of Sensor and Actuator Networks*, 12(2), 28. <https://doi.org/10.3390/jsan12020028>

Dumont, M., Brun, G., Oger, O., Taugourdeau, O., Boukcim, H., Taylor, J. A., & Tisseyre, B. (2023). A novel approach for field sampling optimization incorporating a generic operational cost constraint. *Conference Proceedings*. https://doi.org/10.3920/978-90-8686-947-3_68

Elahi, M., Afolaranmi, S. O., Lastra, J. L. M., & Garcia, J. A. P. (2023). A comprehensive literature review of the applications of AI techniques through the lifecycle of industrial equipments. *Discover Artificial Intelligence Journal*, 3, 3-43. <https://doi.org/10.1007/s44163-023-00089-x>

Embrapa - Embrapa Meio Ambiente. (2022). Agricultura e efeito estufa. Disponível em: <https://www.cnpma.embrapa.br/projetos/index.php3?sec=agrog:::85>. Acesso em: 13 jun. 2022

Emmi, L., Fernandez, R., & Guerrero, J. M. (2023). Editorial: Robotics for smart farms. *Frontiers in Robotics and AI*, 9, 1113440. <https://doi.org/10.3389/frobt.2022.1113440>

Ezziyyani, M., Hamdache, A., Cherrat, L., Laksiby, I., Ezziyyani, M., Alilou, H., Mauri,

J., Chaik, A., & Gilabert, C. E. (2023). Contribution to the development of a technological platform for analysis in precision agriculture for the biovigilance of cryptogamic diseases in strawberry (*Fragaria × ananassa*). Conference Paper. https://doi.org/10.1007/978-3-031-35248-5_57

Fakhar, M. I., & Khalid, M. N. (2023). Satellites to agricultural fields: The role of remote sensing in precision agriculture. *Biological and Agricultural Sciences Research Journal*, 2(14). <https://doi.org/10.54112/basrj.v2023i1.14>

FAO. Food and Agriculture Organization of the United Nations. (2022). Disponível em: <https://www.fao.org/home/en>. Acesso em: 05 nov. 2023

FAO-STAT, Food and Agriculture Organization of the United Nations. (2022). FAOSTAT [Statistical Data]: Major Commodities Importers. Disponível em: https://www.fao.org/faostat/en/#rankings/major_commodities_imports. Acesso em: 21 jun. 2023

Fatima, S. A., Nasim, S. F., & Ahmed, S. (2024). Enhancing agricultural operations: Big data analytics using distributed and parallel computing. ResearchGate. <https://doi.org/10.13140/RG.2.2.25771.21281>

Gamal, M., Elhamahmy, M., Taha, S., & Elmahdy, H. (2024). Improving intrusion detection using LSTM-RNN to protect drones' networks. *Egyptian Informatics Journal*, 2024(100501). <https://doi.org/10.1016/j.eij.2024.100501>

Gangwani, P., Pons, A., Joshi, S., Upadhyay, H., & Lagos, L. (2023). Integration of data science and IoT with blockchain for industry 4.0. *Lecture Notes in Electrical Engineering*, 978-981-19-8730-4, 6. https://doi.org/10.1007/978-981-19-8730-4_6

GAO - U.S. GOVERNMENT ACCOUNTABILITY OFFICE (2024). Precision agriculture: Benefits and challenges for the adoption and use of technology. Disponível em: <https://www.gao.gov/products/gao-24-105962>. Acesso em: 09 ago. 2024

- Gawande, V., Saikanth, D. R. K., Sumithra, B. S., Aravind, S. A., Swamy, G. N., & Chowdhury, M. (2023). Potential of precision farming technologies for eco-friendly agriculture. *International Journal of Plant & Soil Science*, 35(19), 3528. <https://doi.org/10.9734/ijps/2023/v35i193528>
- Gill, R., & Chawla, P. (2021). A review on various techniques to transform traditional farming to precision agriculture. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 2(2), 131–135. <https://turcomat.org/index.php/turkbilm/arte/view/690>
- Gokool, S., Mahomed, M., Kunz, R., Clulow, A., Sibanda, M., Naiken, V., Chetty, K., & Mabhaudhi, T. (2023). Crop monitoring in smallholder farms using unmanned aerial vehicles to facilitate precision agriculture practices: A scoping review and bibliometric analysis. *Sustainability*, 15(4), Artigo 4. <https://doi.org/10.3390/su15043557>
- Gonçalves, P., Pedreiras, P., & Monteiro, A. (2022). Foreword to the special issue on advanced IoT technologies in agriculture. *Applied Sciences*, 12(19). <https://www.mdpi.com/2076-3417/12/19/10102>
- Govi, D., Pappalardo, S., De Marchi, M., & Meggio, F. (2024). From space to field: Combining satellite UAV and agronomic data in an open-source methodology for the validation of NDVI maps in precision viticulture. *Remote Sensing*, 16, 735. <https://doi.org/10.3390/rs16050735>
- Hashemi-Beni, L., Gebrehiwot, A., Karimoddini, A., Shahbazi, A., & Dorbu, F. (2022). Deep convolutional neural networks for weeds and crops discrimination from UAS imagery. *Frontiers in Remote Sensing*, 3, 755939. <https://doi.org/10.3389/frsen.2022.755939>
- He, L. (2023). Variable rate technologies for precision agriculture. In Zhang Q. (eds) *Encyclopedia of Digital Agricultural Technologies*. Springer Cham https://doi.org/10.1007/978-3-030-89123-7_34-3
- He, Q. (2024). Edge computing-oriented smart agricultural supply chain mechanisms with auction and fuzzy neural networks. *Journal of Cloud Computing*, 13. <https://doi.org/10.1186/s13677-024-00626-8>.
- Huerta, M. K., García-Cedeño, A., Guillermo, J. C., & Clotet, R. (2021). Wireless sensor networks applied to precision agriculture: A worldwide literature review with emphasis on Latin America. *IEEE Geoscience and Remote Sensing Magazine*. <https://doi.org/10.1109/MGRS.2020.3044235>
- Huihui, Y., Daoliang, L., & Yingyi, C. (2023). A state-of-the-art review of image motion deblurring techniques in precision agriculture. *Heliyon*, 9(6), e17332. <https://doi.org/10.1016/j.heliyon.2023.e17332>
- Hundal, G. S., Laux, C. M., Buckmaster, D., Sutton, M. J., & Langemeier, M. (2023). Exploring barriers to the adoption of Internet of Things-based precision agriculture practices. *Journal of Agriculture*, 13(163). <https://doi.org/10.3390/agriculture13010163>
- Ijaz, M. F., & Wozniak, M. (2024). Recent advances in big data, machine, and deep learning for precision agriculture. *Frontiers Media SA*. <https://doi.org/10.3389/978-2-8325-4495-2>.
- Javaid, M., Haleen, A., Singh, R. P., Khan, S., & Suman, R. (2021). Blockchain technology applications for Industry 4.0: A literature-based review. *Blockchain: Research and Applications Journal*, 2(4). <https://doi.org/10.1016/j.bcra.2021.100027>
- John, D., Hussin, N., Shahibi, M. S., Ahmad, M., Hashim, H., & Ametefe, D. S. (2023). A systematic review on the factors governing precision agriculture adoption among small-scale farmers. *Outlook on Agriculture*, 52(4), 469–485. <https://doi.org/10.1177/00307270231205640>.
- Johnraja, J., Leelipushpam, P., Shirley, C., & Princess, J. (2024). Impact of cloud computing on the future of smart farming. In *Em Futuristic Trends in Physical Sciences* (pp. 391–420). Springer. https://doi.org/10.1007/978-3-031-51195-0_18.

- Karada, M. S., Bajpai, R., Singh, M., Singh, A. K., Agnihotri, D., & Singh, B. K. (2023). A review on advances in agriculture and agroforestry with GPS and GIS. *International Journal of Plant & Soil Science*, 35(6), 150–160.
<https://doi.org/10.9734/ijpss/2023/v35i62849>
- Karunathilake, E. M. B. M., Le, A. T., Heo, S., Chung, Y. S., & Mansoor, S. (2023). The path to smart farming: Innovations and opportunities in precision agriculture. *Agriculture MDPI*.
<https://doi.org/10.3390/agriculture13081593>.
- Kendall, H., Clark, B., Li, W., Jin, S., Jones, G. D., Chen, J., Taylor, J., Li, Z., & Frewer, L. J. (2022). Precision agriculture technology adoption: A qualitative study of small-scale commercial “family farms” located in the North China Plain. *Precision Agriculture*, 23(1), 319–351.
<https://doi.org/10.1007/s11119-021-09839-2>.
- Kganyago, M., Adjorlolo, C., Mhangara, P., & Tsoeleng, L. (2024). Optical remote sensing of crop biophysical and biochemical parameters: An overview of advances in sensor technologies and machine learning algorithms for precision agriculture. *Computers and Electronics in Agriculture*, 218, 108730.
<https://doi.org/10.1016/j.compag.2024.108730>.
- Kim, M.-Y., & Lee, K. H. (2022). Electrochemical sensors for sustainable precision agriculture—A review. *Frontiers in Chemistry*, 10, 848320.
<https://doi.org/10.3389/fchem.2022.848320>.
- Kim, W.-S. (2023). Data-driven agricultural innovation technology for digital agriculture. *Applied Science Journal*, 13.
<https://doi.org/10.3390/app132011163>.
- Krishnababu, M. E., Rama Devi, B., Soni, A., Panigrahi, C. K., Sudeepthi, B., Rathi, A., & Shukla, A. (2024). A review on precision agriculture navigating the future of farming with AI and IoT. *Asian Journal of Soil Science and Plant Nutrition*, 10(2), 336–349.
<https://doi.org/10.9734/ajsspn/2024/v10i2291>.
- Kumar, I., & Bargavi, D. S. (2024). The role of data science in precision agriculture. *International Research Journal of Modernation in Engineering Technology and Science*, 6.
<https://doi.org/10.56726/IRJMETS50263>.
- Kumar, M., Sinha, A. K., & Kumar, N. (2022). Literature review of smart and precision agriculture by applying IoT module. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 10(IX).
<https://doi.org/10.22214/ijraset.2022.46688>.
- Kumar, R., Farooq, M., & Qureshi, M. (2024b). Advancing precision agriculture through artificial intelligence: Exploring the future of cultivation. *Em Capitulo 10* (pp. 151–165).
<https://doi.org/10.1016/B978-0-443-24001-0.00010-5>.
- Kumar, V., Sharma, K. V., Kedam, N., Patel, A., Kate, T. R., & Rathnayake, U. (2024a). A comprehensive review on smart and sustainable agriculture using IoT technologies. *Agricultural Technology*, 8.
<https://doi.org/10.1016/j.atech.2024.100487>
- Kumari, A., & Bhavanishankar, K. (2023). Precision agriculture. *International Journal of Computing Algorithm*, 12(1).
<https://doi.org/10.20894/IJCOA.101.012.001.004>.
- Kurniawan, D., Indah, D. R., & Sari, P. (2023). Geo-informatics for the future: A systematic literature review on the role of WebGIS in infrastructure planning and development. *Indonesian Journal of Computer Science*.
<https://doi.org/10.33022/ijcs.v12i3.3228>.
- Lamb, J. R., Araújo, E. C., Bazzi, C. L., & Maggi, M. F. (2022). Sampling in precision agriculture: A software for navigation in field activities. *Sustainable Computing: Informatics and Systems Journal*, 36.
<https://doi.org/10.1016/j.suscom.2022.100804>.
- Lanucara, S., Praticò, S., Pioggia, G., Di Fazio, S., & Modica, G. (2024). Web-based spatial decision support system for precision agriculture: A tool for delineating dynamic management unit zones (MUZs). *Agricultural Technology*, 8.
<https://doi.org/10.1016/j.atech.2024.100444>
- Lara, L. M., Gebler, L., Leite Júnior, M. J., & Soares, A. L. (2021). Precision agriculture

- trends in fruit growing from 2016 to 2020. *Revista Brasileira de Fruticultura*, 43(2), e-096. <https://doi.org/10.1590/0100-29452021096>.
- Li, J., Chen, D., Qi, X., Li, Z., Huang, Y., Morris, D., & Tan, X. (2023). Label-efficient learning in agriculture: A comprehensive review. *Computers and Electronics in Agriculture*, 215, 108412. <https://doi.org/10.1016/j.compag.2023.108412>.
- Lu, L., Tian, G., & Hatzenbuehler, P. (2022). How agricultural economists are using big data: A review. *China Agricultural Economic Review*, 14(3), 494–508. <https://doi.org/10.1108/CAER-09-2021-0167>.
- Manjunath, N., Bharatkumar, S. S., Spoorthi, R. B., Chinmayi, N. J., & Tejashwini, C. (2024). Survey on optimal crop prediction using soil and weather analysis. *International Journal of Advanced Research in Science Communication and Technology*, 352–355. <https://doi.org/10.48175/IJARSCT-18154>.
- MAPA - Ministério da Agricultura, Pecuária e Abastecimento do Governo Federal do Brasil. (2022). *Agricultura digital e de precisão*. Disponível em: <https://www.gov.br/agricultura/pt-br/assuntos/sustentabilidade/tecnologia-agropecuaria/agricultura-de-precisao-1>. Acesso em: 13 jul. 2022.
- Maurya, D. K., Maurya, S. K., Kumar, M., Chaubey, C., Gupta, D., Patel, K. K., Mehta, A. K., & Yadav, R. (2024). A review on precision agriculture: An evolution and prospect for the future. *International Journal of Plant & Soil Science*, 36(5), 363–374. <https://doi.org/10.9734/ijpss/2024/v36i54534>.
- Melzer, M., Schulze, L., & Becker, A. (2023). Commercial farm management information systems: A demand-oriented analysis of functions in practical use. *Smart Agricultural Technology*, 4, 100156. <https://doi.org/10.1016/j.atech.2023.100203>
- Memon, K., Umrani, F. A., Baqai, A., & Syed, Z. S. (2023). A review based on comparative analysis of techniques used in precision agriculture. *2023 4th International Conference on Computing Mathematics and Engineering Technologies (iCoMET)*, 1–7. <https://doi.org/10.1109/iCoMET57998.2023.10099182>.
- Mendeley. (2021). Search | Mendeley [Bibliographic reference manager]. Elsevier Company. Disponível em: <https://www.mendeley.com/search/>. Acesso em: 13 jun. 2023.
- Mendes, J., Bueno, L., Oliveira, A., & Gerolamo, M. (2022). Agriculture startups (AgTechs): A bibliometric study. *International Journal of Professional Business Review*, 7, 1–18. <https://doi.org/10.26668/businessreview/2022.v7i2.312>.
- Mendes, J., Pinho, T., Santos, F. N., Sousa, J., Peres, E., Cunha, J. B., Cunha, M., & Morais, R. (2020). Smartphone applications targeting precision agriculture practices—A systematic review. *Agronomy Journal*, 10. <https://doi.org/10.3390/agronomy10060855>
- Messina, G., & Modica, G. (2022). Twenty years of remote sensing applications targeting landscape analysis and environmental issues in olive growing: A review. *Remote Sensing*, 14(21), 5430. <https://doi.org/10.3390/rs14215430>.
- Micheni, E., Machii, J., & Murumba, J. (2022). Internet of things, big data analytics, and deep learning for sustainable precision agriculture. *2022 IST-Africa Conference (IST-Africa)*, 1–12. <https://doi.org/10.23919/IST-Africa56635.2022.9845510>.
- Miguel-Rojas, C., & Pérez-de-Luque, A. (2023). Nanobiosensors and nanoformulations in agriculture: New advances and challenges for sustainable agriculture. *Environmental Technology & Innovation*, 7. <https://doi.org/10.1042/ETLS20230070>.
- Misara, R., Verma, D., Mishra, N., Rai, S. K., & Mishra, S. (2022). Twenty-two years of precision agriculture: A bibliometric review. *Precision Agriculture*, 23(6), Artigo 6. <https://doi.org/10.1007/s11119-022-09969-1>.
- Mitchell, S., Weersink, A., & Bannon, N. (2021). Adoption barriers for precision agriculture technologies in Canadian crop production. *Canadian Journal of Plant Science*,

- 101(3), 412–416. <https://doi.org/10.1139/cjps-2020-0234>.
- Mizik, T. (2023). How can precision farming work on a small scale? A systematic literature review. *Precision Agriculture*, 24, 384–406. <https://doi.org/10.1007/s11119-022-09934-y>
- Muhammed, D., Ahvar, E., Ahvar, S., Trocan, M., Montpetit, M.-J., & Ehsani, R. (2024). Artificial intelligence of things (AIoT) for smart agriculture: A review of architectures, technologies, and solutions. *Journal of Network and Computer Applications*, 228. <https://doi.org/10.1016/j.jnca.2024.103905>.
- Mukati, E. N., Thakur, M. S., Rege, M. P., & Chourasiya, E. A. (2024). Leveraging 5G Drone Technology for Enhanced Agricultural Automation. 8(3), 347–352. www.ijtsrd.com/papers/ijtsrd64894.pdf.
- Musa, P., Sugeru, H., & Wibowo, E. P. (2023). Wireless Sensor Networks for Precision Agriculture: A Review of NPK Sensor Implementations. *Sensors*, 24(1), 51. <https://doi.org/10.3390/s24010051>
- Nagy, S., & Lakatos, V. (2022). The meaning of ICT tools and the control of one in agriculture, with special respect for one for mobile software and applications. *Jelenkőri Társadalmi és Gazdasági Folyamatok*, 17(1–2), 93–108. <https://doi.org/10.14232/jtgf.2022.1-2.93-108>
- NASA - National Aeronautics and Space Administration (2022). WebGIS public - Disponível em: <https://www.nasa.gov/>. Acesso em: 12 jun. 2022.
- Navdeep, Siwach, V., Sehrawat, H., & Dahiya, O. (2022). Hybrid Blockchain Security Mechanism for protection of data related to Precision Agriculture. 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 1–7. <https://doi.org/10.1109/ICRITO56286.2022.9964757>
- Nguyen, L. L. H., Halibas, A., & Quang Nguyen, T. (2023). Determinants of precision agriculture technology adoption in developing countries: A review. *Journal of Crop Improvement*, 37(1), 1–24. <https://doi.org/10.1080/15427528.2022.2080784>
- NitinKamble, N., & Kulkarni, D. A. (2024). INTERNET OF THINGS (IOT) AND SMART WORLD. Em Dr. A. K. Singh, Dr. S. Chapi, Dr. H. K. Pandey, & Mrs. V. Sridhar (Orgs.), *Futuristic Trends in Physical Sciences Volume 3 Book 1* (First, p. 232–265). Iterative International Publishers, Selfpage Developers Pvt Ltd. <https://doi.org/10.58532/V3BKPS1CH13>
- Nyéki, A., & Neményi, M. (2022). Crop Yield Prediction in Precision Agriculture. *Agronomy*, 12(10), 2460. <https://doi.org/10.3390/agronomy12102460>
- Oliveira, M. C., Amaral, P. V., Oliveira, S. F., & Oliveira, J. R. (2023). Improving water efficiency in agriculture using IoT-based systems: A case study. *Environmental Monitoring and Assessment*, 195, 108. <https://doi.org/10.1007/s10661-023-10875-7>.
- Omia, E., Bae, H., Park, E., & Kim, M. S. (2023). Remote Sensing in Field Crop Monitoring: A Comprehensive Review of Sensor Systems, Data Analyses and Recent Advances. *Remote Sensor*. <https://www.mdpi.com/2072-4292/15/2/354>
- Ongadi, P. A. (2024). A comprehensive examination of security and privacy in precision agriculture technologies. *GSC Advanced Research and Reviews*, 18(1), 336–363. <https://doi.org/10.30574/gscarr.2024.18.1.0026>
- Otto, A., Agatz, N., Campbell, J., Golden, B., & Pesch, E. (2018). Optimization approaches for civil applications of unmanned aerial vehicles (UAVs) or aerial drones: A survey. *Networks*, 72(4), 411–458. <https://doi.org/10.1002/net.21818>
- Padhiary, M., Saha, D., Kumar, R., & Sethi, L. N. (2024). Enhancing precision agriculture: A comprehensive review of machine learning and AI vision applications in all-terrain vehicle for farm automation. 8. <https://doi.org/10.1016/j.atech.2024.10048>
- Padhy, S., Alowaidi, M., Dash, S., Alshehri, M., Malla, P. P., Routray, & S., Alhumyani, H. (2023). AgriSecure: A Fog Computing-Based Security Framework for Agriculture 4.0 via

- Blockchain. *Processes*, 11(3), 757. <https://doi.org/10.3390/pr11030757>.
- Padmanabhan, A. A., & Megalingam, R. K. (2022). Autonomous Agricultural Farming Robot for Automatic Ploughing in a Closed Field. 2022 IEEE 19th India Council International Conference (INDICON), 1–6. <https://doi.org/10.1109/INDICON56171.2022.10040212>.
- Pakseresht, A., Yavari, A., Kaliji, S. A., & Hakelius, K. (2023). The intersection of blockchain technology and circular economy in the agri-food sector. *Sustainable Production and Consumption*, 35, 260–274. <https://doi.org/10.1016/j.spc.2022.11.002>.
- Pande, C. B., & Moharir, K. N. (2023). Application of Hyperspectral Remote Sensing Role in Precision Farming and Sustainable Agriculture Under Climate Change: A Review. https://doi.org/10.1007/978-3-031-19059-9_21.
- Pasandideh, F., Da Costa, J. P. J., Kunst, R., Islam, N., Hardjawana, W., & Pignaton De Freitas, E. (2022). A Review of Flying Ad Hoc Networks: Key Characteristics, Applications, and Wireless Technologies. *Remote Sensing*, 14(18), 4459. <https://doi.org/10.3390/rs14184459>.
- Pasquel, D., Roux, S., Richetti, J., Cammarano, D., Tisseyre, B., & Taylor, J. A. (2022). A review of methods to evaluate crop model performance at multiple and changing spatial scales. *Precision Agriculture*, 23(4), Artigo 4. <https://doi.org/10.1007/s11119-022-09885-4>.
- Peppi, L. M., Zauli, M., Manfrini, L., Grappadelli, L. C., Marchi, L. D., & Traverso, P. A. (2023). Low-cost, high-resolution and no-manning distributed sensing system for the continuous monitoring of fruit growth in precision farming. *Acta IMEKO*, 12(2), Artigo 2. <https://doi.org/10.21014/actaimeko.v12i2.1342>.
- Petrović, B., Bumbálek, R., Zoubek, T., Kuneš, R., & Smutný, L. (2024). Application of precision agriculture technologies in Central Europe-review. 15. <https://doi.org/10.1016/j.jafr.2024.101048>.
- Polymeni, S., Plastras, S., Skoutas, D. N., Kormentzas, G., & Skianis, C. (2023). The Impact of 6G-IoT Technologies on the Development of Agriculture 5.0: A Review. *Electronics*, 12(12), 2651. <https://doi.org/10.3390/electronics12122651>.
- PR Newswire, U. G. (2024). Precision Farming and Biotech Innovations Drive Agricultural Growth and Global Food Security. Disponível em: <https://www.prnewswire.com/news-releases/precision-farming-and-biotech-innovations-drive-agricultural-growth-and-global-food-security-302159598.html>. Acesso em: 28 ago. 2024.
- Prey, L., & Schmidhalter, U. (2019). Simulation of satellite reflectance data using high-frequency ground based hyperspectral canopy measurements for in-season estimation of grain yield and grain nitrogen status in winter wheat. 149, 176–187. <https://doi.org/10.1016/j.isprsjprs.2019.01.023>.
- Qiao, Y., Valente, J., Su, D., Zhang, Z., & He, D. (2022). Editorial: AI, sensors and robotics in plant phenotyping and precision agriculture. *Frontiers in Plant Science*, 13, 1064219. <https://doi.org/10.3389/fpls.2022.1064219>.
- Radhika, A., & Masood, Dr. M. S. (2022). Crop Yield Prediction by Integrating Et-DP Dimensionality Reduction and ABP-XGBOOST Technique. *Journal of Internet Services and Information Security*, 12(4), 177–196. <https://doi.org/10.58346/JISIS.2022.I4.013>.
- Radočaj, D., Plaščak, I., & Jurišić, M. (2023a). Global Navigation Satellite Systems as State-of-the-Art Solutions in Precision Agriculture: A Review of Studies Indexed in the Web of Science. *Agriculture*, 13(7), Artigo 7. <https://doi.org/10.3390/agriculture13071417>.
- Radočaj, D., Šiljeg, A., Marinović, R., & Jurišić, M. (2023b). State of Major Vegetation Indices in Precision Agriculture Studies Indexed in Web of Science: A Review. *Agriculture*, 13(3), Artigo 3. <https://doi.org/10.3390/agriculture13030707>.
- Ragazou, K., Garefalakis, A., Zafeiriou, E., & Passas, I. (2022). Agriculture 5.0: A New Strategic Management Mode for a Cut Cost and an Energy Efficient Agriculture Sector.

- Energies, 15(9), 3113. <https://doi.org/10.3390/en15093113>.
- Raghunath, S., & Devi, K. S. (2022). IOT IN PRECISION FARMING FOR A SUSTAINABLE FUTURE. 13(10), 728–764. <https://doi.org/10.47750/pnr.2022.13.S10.078>.
- Rai, N., Pandey, S. K., & Tiwari, K. (2023). Advancements in deep learning techniques for disease detection in precision agriculture: A comprehensive review. *Computers and Electronics in Agriculture*, 190, 106364. <https://doi.org/10.1016/j.compag.2022.106364>.
- Raihan, A. (2024). A Systematic Review of Geographic Information Systems (GIS) in Agriculture for Evidence-Based Decision Making and Sustainability. <https://doi.org/10.56556/gssr.v3i1.636>
- Raouhi, E. M., Lachgar, M., Hrimech, H., & Kartit, A. (2023). Unmanned Aerial Vehicle-based Applications in Smart Farming: A Systematic Review. *International Journal of Advanced Computer Science and Applications*, 14(6). <https://doi.org/10.14569/IJACSA.2023.01406123>.
- Reda, M., Refaai, A., Dattu, V. S., Gireesh, N., Dixit, N., & Sandeep, C. H. (2023). Retracted: Application of IoT-Based Drones in Precision Agriculture for Pest Control. *Advances in Materials Science and Engineering*, 2023(1), 1–1. <https://doi.org/10.1155/2023/9763026>.
- Rehman, A. U., Alamoudi, Y., Khalid, H. M., Morchid, A., Muyeen, S. M., & Abdelaziz, A. Y. (2024). Smart agriculture technology: An integrated framework of renewable energy resources, IoT-based energy management, and precision robotics. 9. <https://doi.org/10.1016/j.cles.2024.100132>.
- Rehman, U., Iqbal, T., Hussain, S., Cheema, M. J. M., Iqbal, F., & Basit, A., (2023). Drone and Robotics Roadmap for Agriculture Crops in Pakistan: A Review. *Environmental Sciences Proceeding Journal*, p. 23-39. <https://doi.org/10.3390/environsciproc2022023039>.
- Rejeb, A., Abdollahi, A., Rejeb, K., & Treiblmaier, H. (2022). Drones in agriculture: A review and bibliometric analysis. *Computers and Electronics in Agriculture*, 198, 107017. <https://doi.org/10.1016/j.compag.2022.107017>.
- Rennings, M., Baaden, P., Block, C., John, M., & Bröring, S. (2024). Assessing emerging sustainability-oriented technologies: The case of precision agriculture. *Scientometrics*, 129(6), 2969–2998. <https://doi.org/10.1007/s11192-024-05022-2>.
- Rimpika, Anushi, Manasa, S., Anusha K. N., Sharma, S., Thakur, A., Shilpa, & Sood, A. (2023). An Overview of Precision Farming. *International Journal of Environment and Climate Change*, 13(12), 441–456. <https://doi.org/10.9734/ijecc/2023/v13i123701>.
- Rivera, G., Porras, R., Florencia, R., & Sánchez-Solís, J. P. (2023). LiDAR applications in precision agriculture for cultivating crops: A review of recent advances. *Computers and Electronics in Agriculture*, 207, 107737. <https://doi.org/10.1016/j.compag.2023.107737>.
- Sanaeifar, A., Guindo, M. L., Bakhshipour, A., Fazayeli, H., Li, X., & Yang, C. (2023). Advancing precision agriculture: The potential of deep learning for cereal plant head detection. 209. <https://doi.org/10.1016/j.compag.2023.107875>.
- Sanyaolu, M., & Sadowski, A. (2024). The Role of Precision Agriculture Technologies in Enhancing Sustainable Agriculture. *Sustainability*, 16(15), 6668. <https://doi.org/10.3390/su16156668>.
- Saranya, T., Deisy, C., Sridevi, S., & Anbananthen, K. S. M. (2023). A comparative study of deep learning and Internet of Things for precision agriculture. *Engineering Applications of Artificial Intelligence*, 122, 106034. <https://doi.org/10.1016/j.engappai.2023.106034>.
- Sathya, D., Thangamani, R. B., & Balaji, S., (2024). The Revolution of Edge Computing in Smart Farming. *Intelligent Robots and Drones for Precision Agriculture Book*, Springer Link, p.351-389, https://doi.org/10.1007/978-3-031-51195-0_17.
- Shahi, T. B., Xu, C.-Y., Neupane, A., Guo, W., Shahi, T. B., Xu, C.-Y., Neupane, A., & Guo,

- W. (2023). Machine learning methods for precision agriculture with UAV imagery: A review. *Electronic Research Archive*, 30(era-30-12-218), Artigo era-30-12-218. <https://doi.org/10.3934/era.2022218>.
- Sharma, K., & Shivandu, S. K. (2024). Integrating artificial intelligence and Internet of Things (IoT) for enhanced crop monitoring and management in precision agriculture. 5. <https://doi.org/10.1016/j.sintl.2024.100292>
- Shukla, B. K., Maurya, N., & Sharma, M. (2023). Advancements in Sensor-Based Technologies for Precision Agriculture: An Exploration of Interoperability, Analytics and Deployment Strategies. *ECSA 2023*, 22. <https://doi.org/10.3390/ecsa-10-16051>.
- Siddegowda, C.J., & Devi, A.J., (2021). A Study on the Role of Precision Agriculture in Agro-Industry. *International Journal of Applied Engineering and Management Letters*, ISSN: 2581-7000, vol. 5, n. 2, <https://doi.org/10.47992/IJAEML.2581.7000.0103>.
- Singh, N., Kapoor, S. (2024). Agtech platforms: Complementors and value propositions. *Technology Analysis & Strategic Management*, 1–16. <https://doi.org/10.1080/09537325.2024.2306636>.
- Singh, R., Aernouts, M., De Meyer, M., Weyn, M., & Berkvens, R. (2020). Leveraging LoRaWAN Technology for Precision Agriculture in Greenhouses. *Sensors*, 20(7), 1827. <https://doi.org/10.3390/s20071827>.
- Sivakumar, M., & Tyj, N. M. (2021). A Literature Survey of Unmanned Aerial Vehicle Usage for Civil Applications. https://www.scielo.br/j/jatm/a/vnWkfk66h5Vv_sRhn6CxjLKx/?lang=en.
- Sivasangari, A., Teja, A. K., S, G., Ajitha, P., Gomathi, R. M., & Vignesh. (2023). Revolutionizing Agriculture: Developing Autonomous Robots for Precise Farming. 2023 International Conference on Inventive Computation Technologies (ICICT), 1461–1468. <https://doi.org/10.1109/ICICT57646.2023.10134507>.
- Sott, M. K., Nascimento, L. D. S., Foguesatto, C. R., Furstenau, L. B., Faccin, K., Zawislak, P. A., Mellado, B., Kong, J. D., & Bragazzi, N. L. (2021). A Bibliometric Network Analysis of Recent Publications on Digital Agriculture to Depict Strategic Themes and Evolution Structure. *Sensors*, 21(23), 7889. <https://doi.org/10.3390/s21237889>.
- Sourav, A. I., & Emanuel, A. W. R. (2021). Recent Trends of Big Data in Precision Agriculture: A Review. *IOP Conference Series: Materials Science and Engineering*, 1096(1), 012081. <https://doi.org/10.1088/1757-899X/1096/1/012081>.
- Sousa, A. S. de, Oliveira, G. S. de, & Alves, L. H. (2021). A Pesquisa Bibliográfica: Princípios E Fundamentos. *Cadernos da FUCAMP*, 20(43), Artigo 43. <https://revistas.fucamp.edu.br/index.php/cadernos/article/view/2336>.
- Soussi, A., Zero, E., Sacile, R., Trincherro, D., & Fossa, M. (2024). Smart Sensors and Smart Data for Precision Agriculture: A Review. *Sensors*, 24(8), 2647. <https://doi.org/10.3390/s24082647>.
- Sun, C., Zhou, J., Ma, Y., Xu, Y., Pan, B., & Zhang, Z. (2022). Frontiers | A review of remote sensing for potato traits characterization in precision agriculture. <https://doi.org/10.3389/fpls.2022.871859>
- Szira, Z., Varga, E., Csegódi, T. L., & Milics, G. (2023). The Benefits, Challenges and Legal Regulation of Precision Farming in the European Union. *EU Agrarian Law*, 12(1), 1–7. <https://doi.org/10.2478/eual-2023-0001>.
- Talaat, F. M. (2023). Crop yield prediction algorithm (CYPA) in precision agriculture based on IoT techniques and climate changes. *Neural Computing and Applications*, 35(23), 17281–17292. <https://doi.org/10.1007/s00521-023-08619-5>.
- Tomaszewski, L., Kołakowski, R., & Zagórd, M., (2022). Application of Mobile Networks (5G and beyond) in Precision Agriculture. In: Maglogiannis, I., Iliadis, L., Macintyre, J., Cortez, P. (eds) *Artificial Intelligence Applications and Innovations. AIAI 2022 IFIP WG 12.5 International Workshops. AIAI 2022*.

- IFIP Advances in Information and Communication Technology, vol 652. Springer, Cham, https://doi.org/10.1007/978-3-031-08341-9_7.
- Tyagi, R., & Pandey, P. C. (2024). Applications of drones in precision agriculture: Future of smart and sustainable farming (p. 429–453). <https://doi.org/10.1016/B978-0-323-91068-2.00003-5>.
- U-Minn, Tallaksen, J., & Minneapolis Star, T. (2022). Project Manager Information. <https://www.precisionfarmingdealer.com/articles/4952-umn-unveils-plans-for-220-million-precision-ag-research-complex>. <https://www.minnesotaalumni.org/stories/measured-growth>.
- Ünal, İ., Kabaş, Ö., Eceoğlu, O., & Moiceanu, G. (2023). Adaptive Multi-Robot Communication System and Collision Avoidance Algorithm for Precision Agriculture. *Applied Sciences*, 13(15), 8602. <https://doi.org/10.3390/app13158602>.
- Upadhyay, A., Zhang, Y., Koparan, C., Rai, N., Howatt, K., & Bajwa, S. (2024). Advances in ground robotic technologies for site-specific weed management in precision agriculture: A review. 224. <https://doi.org/10.1016/j.compag.2024.109363>.
- USDA, U.S. Department of Agriculture, (2022) – Agricultural Research Service. List of government entities in research and resources for broadband access in the United States. - Disponível em: <https://fdc.nal.usda.gov/>. Acesso em: 12 jun. 2022.
- Verma, A. (2021). SVM, CNN and VGG16 Classifiers of Artificial Intelligence Used for the Detection of Diseases of Rice Crop: A Review. https://doi.org/10.1007/978-981-16-5157-1_71.
- Vinod, C. S. S., Anand, H. S., & Albaaji, G. F. (2024). Precision farming for sustainability: An agricultural intelligence model. <https://www.sciencedirect.com/science/article/abs/pii/S0168169924007774>. Vol.226. <https://doi.org/10.1016/j.compag.2024.109386>
- Vishnu, M., Rajesh, G. M., Damor, J. S., Pandey, S. K., Kumar, P., & Singh, B. (2024). Remote Sensing and Geographic Information Systems for Precision Agriculture: A Review. *International Journal of Environment and Climate Change*. <https://doi.org/10.9734/ijecc/2024/v14i23945>.
- Vrchota, J., Pech, M., & Švepešová, I. (2022). Precision Agriculture Technologies for Crop and Livestock Production in the Czech Republic. *Agriculture*, 12(8), 1080. <https://doi.org/10.3390/agriculture12081080>.
- Wang, X.-H., Wen, Y.-H., Li, P.-Y., Li, Y.-X., Qin, H.-Y., Zhao, J.-X., Zhai, X.-J., Yang, W.-C., & Wu, L.-D. (2024). Flexible sensors for precision agriculture: A mini review. 180, 117946. <https://doi.org/10.1016/j.trac.2024.117946>.
- Wei, Z., Ning, F., Jiang, K., Wang, Y., Bi, Z., Duan, Q., Zhang, J., & Li, R. (2024). CatBoost-based Intrusion Detection Method for the Physical Layer of Smart Agriculture. *ITM Web of Conferences*, 60, 00009. <https://doi.org/10.1051/itmconf/20246000009>.
- Wrzecińska, M., Czerniawska-Piątkowska, E., Kowalewska, I., Kowalczyk, A., Mylostyvyi, R., & Stefaniak, W. (2023). Agriculture in the face of new digitization technologies. *UKRAINIAN BLACK SEA REGION AGRARIAN SCIENCE*, 27(3), 9–17. <https://doi.org/10.56407/bs.agrarian/3.2023.09>.
- Yazdinejad, A., Zolfaghari, B., Azmoodeh, A., Dehghantanha, A., Karimipour, H., Fraser, E., Green, A. G., Russell, C., & Duncan, E. (2021). A Review on Security of Smart Farming and Precision Agriculture: Security Aspects, Attacks, Threats and Countermeasures. *Applied Sciences*, 11(16), 7518. <https://doi.org/10.3390/app11167518>.
- Zain, M., Ma, H., Rahman, S. U., Nuruzzaman, Md., Chaudhary, S., Azeem, I., Mehmood, F., Duan, A., & Sun, C. (2024). Nanotechnology in precision agriculture: Advancing towards sustainable crop production. 206. <https://doi.org/10.1016/j.plaphy.2023.108244>.
- Zaman, Q. U., (2023). Chapter 1 - Precision agriculture technology: a pathway toward sustainable agriculture. *Precision Agriculture Journal*, Elsevier, Evolution, Insights and Emerging Trends, p.1-17, <https://doi.org/10.1016/B978-0-443-18953-1.00013-1>.

Zanin, A. R. A. (2022). Reduction of pesticide application via real-time precision spraying. *Sci Rep* 12, 5638 (2022). <https://doi.org/10.1038/s41598-022-09607-w>.

Zhao, W., Li, T., Qi, B., Nie, Q., & Runge, T. (2021). Terrain Analytics for Precision Agriculture with Automated Vehicle Sensors and Data Fusion. *Sustainability*, 13(5), 2905. <https://doi.org/10.3390/su13052905>.

Zhou, Z., Majeed, Y., Diverres Naranjo, G., & Gambacorta, E. M. T. (2021). Assessment for crop water stress with infrared thermal imagery in precision agriculture: A review and future prospects for deep learning applications. *Computers and Electronics in Agriculture*, 182, 106019. <https://doi.org/10.1016/j.compag.2021.106019>.