

NATHAN LAMOUNIER LIMA

**ESTIMATIVA DA ALTURA DO DOSSEL E PREDIÇÃO DO ACAMAMENTO
EM PARCELAS DE SOJA VIA ANÁLISE DE IMAGENS**

Tese apresentada à Universidade Federal de Viçosa,
como parte das exigências do Programa de Pós-
Graduação em Fitotecnia, para obtenção do título de
Doctor Scientiae.

Orientador: Felipe Lopes da Silva

Coorientadores: Lucas de Paula Corrêdo
Luiz Alexandre Peternelli
Domingos Sárvio Magalhães Valente

VIÇOSA - MINAS GERAIS

2024

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

T

L732e
2024

Lima, Nathan Lamounier, 1992-

Estimativa da altura do dossel e predição do acamamento
em parcelas de soja via análise de imagens / Nathan Lamounier
Lima. – Viçosa, MG, 2024.

1 tese eletrônica (90 f.): il. (algumas color.).

Inclui apêndices.

Orientador: Felipe Lopes da Silva.

Tese (doutorado) - Universidade Federal de Viçosa,
Departamento de Agronomia, 2024.

Referências bibliográficas: f. 70-80.

DOI: <https://doi.org/10.47328/ufvbbt.2024.643>

Modo de acesso: World Wide Web.

1. *Glycine max* - Tamanho - Métodos estatísticos. 2.
Fotogrametria. 3. Análise de séries temporais. 4. Análise de
regressão logística. 5. Árvores (Teoria dos grafos). I. Silva,
Felipe Lopes da, 1981-. II. Universidade Federal de Viçosa.
Departamento de Agronomia. Programa de Pós-Graduação em
Fitotecnia. III. Título.

CDD 22. ed. 633.34

NATHAN LAMOUNIER LIMA

**ESTIMATIVA DA ALTURA DO DOSSEL E PREDIÇÃO DO ACAMAMENTO
EM PARCELAS DE SOJA VIA ANÁLISE DE IMAGENS**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Fitotecnia, para obtenção do título de *Doctor Scientiae*.

APROVADA: 12 de julho de 2024.

Assentimento:

Nathan Lamounier Lima

Autor

Felipe Lopes da Silva

Orientador

AGRADECIMENTOS

Aos meus familiares, em especial à minha mãe, Zilka Lamounier Alves de Lima, e ao meu pai, Gilberto de Araújo Lima, pelo apoio incondicional e confiança depositada em mim ao longo desta jornada.

À minha companheira, Brunna Figueiredo Duarte, pela companhia constante e pelo suporte durante toda a minha trajetória acadêmica.

À Universidade Federal de Viçosa, pela oportunidade de realizar a pós-graduação e pelo ambiente de crescimento acadêmico e pessoal.

Ao meu orientador, Professor Felipe Lopes da Silva, pela paciência, sabedoria e conhecimentos compartilhados.

Aos colegas do Programa de Melhoramento da Soja da Universidade Federal de Viçosa, pelo suporte nas atividades de pesquisa e pelas ricas experiências compartilhadas.

Ao corpo docente do Programa de Pós-Graduação em Fitotecnia, pelos ensinamentos e pela dedicação constante ao desenvolvimento de nossos conhecimentos.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), pela concessão da bolsa de estudos e pelo suporte financeiro indispensável para a realização desta pesquisa.

Este trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

RESUMO

LIMA, Nathan Lamounier, D.Sc., Universidade Federal de Viçosa, julho de 2024. **Estimativa da altura do dossel e predição do acamamento em parcelas de soja via análise de imagens.** Orientador: Felipe Lopes da Silva. Coorientadores: Lucas de Paula Corrêdo, Luiz Alexandre Peternelli e Domingos Sárvio Magalhães Valente.

As avaliações da altura de plantas e da ocorrência e severidade do acamamento em parcelas de soja são essenciais para programas de melhoramento. No entanto, a avaliação tradicional dessas características é laboriosa e muitas vezes baseada em escalas visuais. O sensoriamento remoto via veículos aéreos não tripulados (VANT) é uma alternativa precisa e eficiente para avaliações fenotípicas. Modelos de Random Forest (MRF) são algoritmos de aprendizado de máquina que possibilitam a classificação de parcelas experimentais com base em padrões em variáveis preditivas. Os objetivos do trabalho foram avaliar o potencial do uso de modelos digitais de elevação, estimados via análise de imagens, na estimação da altura do dossel nos estádios de florescimento e maturação, e avaliar o potencial de MRF em detectar e classificar parcelas de soja quanto à severidade do acamamento através de séries temporais de alturas do dossel. Dois ensaios com um total de 307 genótipos de soja foram avaliados quanto ao número de dias até o florescimento (DIASF) e até a maturação (DIASM), comprimento de plantas no florescimento (COMPF) e na maturação (COMPM) e nível de acamamento (AC). O AC foi estimado por uma escala de níveis correspondentes ao ângulo de inclinação da haste principal. Imagens aéreas foram adquiridas através de 19 voos com um VANT com sensor RGB, entre o florescimento e a maturação. Modelos digitais de altura do dossel foram estimados, e a altura do dossel estimada (ALTDI) foi determinada pelos percentis 50, 75, 90, 95, 99 e pelo intervalo de altura predominante. As ALTDI de voos próximos ao DIASF e ao DIASM foram comparadas com COMPF e com a altura do dossel calculada a partir de COMPM e AC, respectivamente. A acurácia das ALTDI foi mensurada pela raiz do erro quadrático médio (RMSE) e raiz do erro quadrático médio percentual (RMSE%). Séries temporais foram construídas com base nas ALTDIs de cada parcela e curvas logísticas foram ajustadas às séries temporais. Os parâmetros dessas curvas foram usados para treinar os MRFs para a classificação de AC, a partir de cada método de estimação de ALTDI (ME_ALTDI). Os MRFs foram avaliados através da acurácia balanceada, F1-Score e coeficiente Kappa. O percentil 50 estimou a altura do dossel no florescimento com acurácia satisfatória (RMSE = 12,9 e RMSE% = 13,5%). O percentil

99 foi superior na estimação da altura do dossel na maturação (RMSE = 28,8 e RMSE% = 18,7%), mas houve um comprometimento da acurácia da ALTDI em parcelas com plantas predominantemente eretas devido à redução da superfície das plantas causada pela senescência e queda de folhas causada pela maturação. Os MRFs treinados a partir do percentil 50 mostraram maior potencial na classificação de AC (acurácia balanceada = 0,5494, F1-Score = 0,2360 e coeficiente Kappa = 0,0876) e na detecção do acamamento (acurácia balanceada = 0,5833, F1-Score = 0,5740 e coeficiente Kappa = 0,1630), mas, devido à baixa acurácia das ALTDI próximas à maturação, não mostraram desempenho suficiente para justificar sua recomendação. A metodologia pode ter maior sucesso com sensores que representem as alturas das plantas de forma mais acurada e detalhada.

Palavras-chave: *Glycine max*. RGB. *Surface from Motion*. Série temporal. Curva logística. Random Forest.

ABSTRACT

LIMA, Nathan Lamounier, D.Sc., Universidade Federal de Viçosa, July, 2024.
Estimation of canopy height and lodging prediction in soybean plots through image analysis. Adviser: Felipe Lopes da Silva. Co-advisers: Lucas de Paula Corrêdo, Luiz Alexandre Peternelli e Domingos Sárvio Magalhães Valente.

Assessing plant height and the occurrence and severity of lodging in soybean plots is essential for breeding programs. However, traditional assessment methods are labor-intensive, and often based on visual scales. Remote sensing via unmanned aerial vehicles (UAVs) offers an accurate and efficient alternative for phenotypic evaluations. Random forest models (RFM) are machine learning algorithms that enable the classification of experimental plots based on patterns in predictive variables. The objective of this study was to evaluate the potential of using digital elevation models, estimated through image analysis, for canopy height estimation at flowering and maturity stages, and to assess the potential of RFM in detecting lodging and classifying its severity in soybean plots through temporal series of canopy heights. Two trials, comprising a total of 307 soybean genotypes, were evaluated for days to flowering (DAYSF), days to maturation (DAYSM), plant length at flowering (LENGF), plant length at maturity (LENGM), and lodging severity (LS). LS was estimated on a scale of levels corresponding to the angle of inclination of the main stem. Aerial images were acquired through 19 UAV flights with an RGB sensor between flowering and maturity. Canopy height models were estimated, and the estimated canopy height (ECH) was determined using the 50th, 75th, 90th, 95th, 99th percentiles, and the predominant height range. ECH from flights executed close to DAYSF and DAYSM were compared with LF and canopy height calculated from LENGM and LS, respectively. ECH accuracy was measured using the root mean square error (RMSE) and RMSE%. Temporal series were constructed based on the ECH, and logistic curves were fitted to the series. Parameters from these curves were used to train RFM for LS classification, using each ALTDI estimator. RFMs were evaluated using balanced accuracy, F1-Score, and Kappa coefficient. The 50th percentile accurately estimated canopy height at flowering (RMSE = 12.9 and RMSE% = 13.5%). The 99th percentile was superior in estimating canopy height at maturity (RMSE = 28.8 and RMSE% = 18.7%), but ECH accuracy in plots with predominantly erect plants was compromised due to reduced plant surface area caused by leaf senescence and fall at maturity. RFM trained using the 50th percentile showed greater potential in LS

classification (balanced accuracy = 0.5494, F1-Score = 0.2360, and Kappa = 0.0876) and lodging detection (balanced accuracy = 0.5833, F1-Score = 0.5740, and Kappa = 0.1630), but due to low ECH accuracy near maturity, they did not perform sufficiently well to justify their recommendation. The methodology may be more successful with sensors that are able to describe plant heights with more details and more accurately.

Keywords: *Glycine max*. RGB. Surface from Motion. Temporal series. Logistic curve. Random Forest.

LISTA DE SIGLAS E ABREVIATURAS

AC	nível de acamamento
AC Predito	nível de acamamento predito através do modelo de Random Forest
AC Real	nível de acamamento mensurado no campo
ALTDC	altura do dossel avaliada no campo
ALTDI	altura do dossel estimadas por análise de imagens
b	limite inferior da curva ajustada
cm	centímetro
COMPF	comprimento de plantas no florescimento
COMPM	comprimento de plantas na maturação
DAP	dias após o plantio
DIASF	número de dias até o florescimento
DIASM	número de dias até a maturação
DIASV	dia relativo da execução do voo
E1	Ensaio 1
E2	Ensaio 2
ExG	índice de vegetação Excesso de Verde normalizado
IAP	intervalo de altura predominante
k	inclinação máxima
L	deslocamento máximo da curva ajustada em relação a seu limite inferior
LiDAR	Light Detection and Ranging
m	metro
MDE	modelos digitais de elevação
ME_ALTDI	método de estimação da altura do dossel por análise de imagens
MRF	modelo de Random Forest
MRF_IAP	modelo de Random Forest treinado com parâmetros da curva logística ajustada à série temporal de alturas estimadas através do intervalo de altura predominante
MRF_P50	modelo de Random Forest treinado com parâmetros da curva logística ajustada à série temporal de alturas estimadas através do percentil 50%
MRF_P75	modelo de Random Forest treinado com parâmetros da curva logística ajustada à série temporal de alturas estimadas através do percentil 75%

MRF_P90	modelo de Random Forest treinado com parâmetros da curva logística ajustada à série temporal de alturas estimadas através do percentil 90%
MRF_P95	modelo de Random Forest treinado com parâmetros da curva logística ajustada à série temporal de alturas estimadas através do percentil 95%
MRF_P99	modelo de Random Forest treinado com parâmetros da curva logística ajustada à série temporal de alturas estimadas através do percentil 99%
N_ALTDI	alturas do dossel normalizadas pelo valor máximo
P50	percentil 50%
P75	percentil 75%
P90	percentil 90%
P95	percentil 95%
P99	percentil 99%
PCT	ponto de controle terrestre
R2	coeficiente de determinação
RMSE	raiz do erro quadrático médio
RMSE%	raiz do erro quadrático médio em porcentagem
SfM	<i>Surface from motion</i>
VANT	veículos aéreos não tripulados
x0	ponto de inflexão no eixo x

SUMÁRIO

1. INTRODUÇÃO.....	11
2. MATERIAL E MÉTODOS.....	17
2.1. Material genético	17
2.2. Execução experimental	17
2.3. Avaliações no campo	18
2.3.1. Características avaliadas	18
2.3.2. Análise estatística das características avaliadas no campo	20
2.4. Análise de imagens aéreas	21
2.4.1. Aquisição de imagens aéreas	21
2.4.2. Processamento de imagens	22
2.5. Estimação da altura do dossel via análise de imagens	25
2.5.1. Avaliação das alturas do dossel via análise de imagens	25
2.6. Estimação do acamamento via análise de imagens	27
2.6.1. Pré-processamento de dados	27
2.6.2. Ajuste da curva logística.....	27
2.6.3. Modelos de Random Forest	28
3. RESULTADOS E DISCUSSÃO	31
3.1. Avaliações no campo	31
3.1.1. Análise estatística das características avaliadas no campo	34
3.2. Estimação da altura do dossel via análise de imagens.....	35
3.3. Estimação do acamamento via análise de imagens	44
3.3.1. Série temporal de alturas do dossel normalizadas pelo valor máximo	44
3.3.2. Ajuste da curva logística.....	47
3.3.3. Modelos de Random Forest	50
4. CONCLUSÕES	69
REFERÊNCIAS	70
APÊNDICE	81

1. INTRODUÇÃO

A soja (*Glycine max* (L) Merrill) desempenha um papel crucial na agricultura global devido à sua importância nutricional, econômica e agrônômica. Seus grãos, ricos em proteínas, óleos e vitaminas essenciais, são utilizados na alimentação humana e na produção de rações animais. Além disso, seus derivados são utilizados em diversos segmentos da indústria (Pagano & Miransari, 2016; Triandita & Putri, 2019).

A produção e exportação de soja desempenham um papel central na economia brasileira, com contribuições importantes para a balança comercial e o desenvolvimento econômico (Rabelo et al., 2023). A produtividade da cultura no país mostra uma tendência crescente consistente ao longo do tempo. Avanços no manejo, proteção de plantas, adubação, desenvolvimento de novos cultivares e a digitalização do campo e das tomadas de decisão dos agricultores são alguns dos fatores que contribuem para o aumento observado da produção (Pereira & Brisola, 2022).

Apesar disso, a produtividade de lavouras de soja pode ser limitada por fatores biológicos, como a incidência de nematoides e de doenças foliares e vasculares (Lin et al., 2022), fatores ambientais, como altas temperaturas, secas e irregularidades na precipitação durante o período de cultivo (Cober & Morrison, 2022). Além disso, a interação entre fatores biológicos, ambientais e relacionados ao posicionamento incorreto de cultivares pode limitar a expressão do potencial genético e reduzir a adaptabilidade e estabilidade dos genótipos cultivados (Kehl et al., 2022).

O acamamento de plantas é caracterizado pela inclinação das hastes das plantas em direção ao solo. Na cultura da soja, esse fenômeno pode gerar reduções de produtividade, devido, principalmente, à dificuldade da colheita mecânica, ao aumento da incidência de doenças e à redução da qualidade dos grãos de vagens em contato com o solo (Tomohiro & Homma, 2023). A incidência e severidade do acamamento podem ser potencializados por fatores ambientais, como a incidência de chuvas e ventos intensos (Wu et al., 2022) e fatores culturais, como altas densidades de plantio (Xu et al., 2020).

A suscetibilidade ao acamamento é uma característica poligênica, de herança complexa e correlacionada a várias outras características (Kato et al., 2021). Estudos de mapeamento de QTL encontraram regiões que explicam entre 2,21 e 20,17% da variação fenotípica observada (Sun et al., 2022). A altura de plantas e o número de dias até a maturação são as principais características agrônômicas que influenciam a susceptibilidade de genótipos de soja ao acamamento, e apresentam correlações positivas

com magnitude moderada (Lee et al., 2015). Além disso, o número de ramificações, comprimento do internó e número de nós da planta de soja também apresentam correlações genótípicas positivas com a incidência do acamamento, enquanto o peso das raízes, e o diâmetro e resistência da haste principal são características que apresentam correlações genótípicas negativas (Chen et al., 2017).

A seleção de genótipos de soja resistentes ao acamamento é imprescindível para o desenvolvimento de cultivares que suportem maiores produtividades e tolerem eventos climáticos extremos sem prejuízos à qualidade dos grãos (Kato et al., 2021; Liu et al., 2017). Apesar disso, a avaliação fenotípica tradicional do acamamento em parcelas experimentais de soja se baseia em escalas de notas visuais, o que demanda intensa mão de obra especializada e a torna subjetiva, uma vez que não se baseia em mensurações quantificáveis (Bai et al., 2022; Lee et al., 2015).

A adoção de veículos aéreos não tripulados (VANT) no monitoramento de lavouras via sensoriamento tem se destacado no cenário agrícola devido à sua eficiência e praticidade (Ibiev et al., 2022). As altas resoluções espaciais (capacidade de detecção e identificação de objetos com alto detalhamento) e temporais (possibilidade de realizar avaliações sequenciais e sucessivas em curtos períodos de tempo) alcançáveis pelo uso dos VANTs no sensoriamento remoto permite monitoramento de grandes áreas com precisões e frequências previamente impraticáveis. O monitoramento da sanidade das lavouras, detecção de ervas daninhas e estimação de características agronômicas, como o índice de área foliar, altura de plantas, conteúdo de clorofila e produtividade estão entre os usos mais comuns da análise de imagens aéreas na agricultura estão o (Hassani et al., 2023; Radočaj et al., 2023; Shu et al., 2022).

A funcionalidade dos VANTs depende dos sensores acoplados a eles. Os principais sensores utilizados na agricultura são os RGB, que têm baixo custo e capturam imagens nas faixas espectrais do vermelho, verde e azul; multiespectrais, que têm maior amplitude espectral e operam nas faixas dos sensores RGB e no infravermelho próximo; sensores térmicos, que detectam radiação infravermelha e a converte em informações de temperatura; e sensores LiDAR (Light Detection and Ranging), que medem a distância e determinam o formato de objetos através do tempo entre emissão e retorno de pulsos de lasers (Radočaj et al., 2023).

A fenotipagem de alto rendimento através da análise de imagens aéreas é uma metodologia inovadora que utiliza sensores, visão computacional, e algoritmos de aprendizado de máquinas para avaliar características agronômicas, morfológicas e

fisiológicas de forma rápida e precisa (Anand & Prashant, 2023). Devido à sua alta eficiência, aplicabilidade em grandes áreas e reduzida necessidade de mão de obra, essa metodologia têm sido amplamente adotada em programas de melhoramento genético e têm mostrado grande potencial em aumentar a acurácia seletiva e reduzir o tempo de desenvolvimento de novas cultivares (Bhandari et al., 2023; Herr et al., 2023). O cálculo de índices de vegetação, baseados nos valores das faixas espectrais detectadas pelos sensores, permite a avaliação de forma indireta de diferentes características relacionadas à sanidade, vigor, resposta a estresses ambientais e potencial produtivo de diversas culturas. Além disso, algoritmos de aprendizado de máquinas e técnicas fotogramétricas podem ser utilizadas durante e após o processamento de imagens para gerar informações relacionadas à altura, formato, volume e textura de parcelas experimentais (Abebe et al., 2023; Buelvas et al., 2023).

A *Surface from motion* (SfM) é uma técnica de fotogrametria capaz gerar nuvens de pontos em três dimensões através de imagens sobrepostas. As nuvens de pontos, por sua vez, permitem a reconstrução de modelos digitais de elevação (MDE) que representam a superfície tridimensional do cenário capturado. Essa técnica pode ser utilizada com imagens aéreas de diferentes sensores, inclusive dos sensores RGB, e é uma alternativa de melhor relação custo-eficiência ao mapeamento topográfico por escaneamento a laser (Fan, 2020). Os MDEs gerados por SfM têm sido empregados no monitoramento ambiental e de áreas agrícolas para a estimação da biomassa, altura do dossel e para a detecção de áreas com plantas acamadas (Acorsi et al., 2019; Grüner et al., 2019).

A altura do dossel e a altura de plantas são características correlacionadas, comumente avaliada em programas de melhoramento genético de soja devido à sua correlação direta com a produtividade de grãos e desempenho geral dos cultivares. Além disso, essas características são diretamente correlacionadas à maior susceptibilidade ao tombamento de plantas e, por isso, seu monitoramento em populações de melhoramento genético é essencial para manter o equilíbrio entre ganhos de produtividade e resistência ao acamamento (Bhat et al., 2022; Chen et al., 2017).

Existem exemplos da aplicação de MDEs gerados via SfM na estimação da altura do dossel de plantas com diferentes níveis de sucesso em parcelas de soja (Borra-Serrano et al., 2020; Maimaitijiang et al., 2019), milho (Acorsi et al., 2019; Chu et al., 2018; Malambo et al., 2018), sorgo (Chang et al., 2017; Han et al., 2018; Malambo et al., 2018), trigo-sarraceno (Murakami et al., 2012) e gramíneas (Obanawa et al., 2020).

Uma vez que o tombamento leva à redução da altura das plantas em relação ao nível do solo, a detecção de reduções inesperadas na altura do dossel de parcelas experimentais permite inferir sobre a incidência do acamamento. Estimativas sequenciais da altura do dossel permitem a construção de curvas de crescimento das plantas ao longo de seu ciclo de desenvolvimento (Wang et al., 2020). O estudo de curvas de crescimento definidas como séries temporais da altura do dossel em ensaios de melhoramento genético apresenta grande potencial na detecção e quantificação da severidade do acamamento (Wilke et al., 2019).

Pugh et al., (2018) utilizaram séries temporais de altura de plantas estimadas a partir de imagens aéreas e SfM para desenvolver curvas de crescimento e identificar variações genotípicas em parcelas de milho e sorgo. Os autores destacaram o potencial dessa abordagem devido à alta resolução temporal obtida por meio da realização da estimação sequencial da altura de plantas, durante todo o ciclo de desenvolvimento. A utilização de séries temporais de imagens aéreas permite a construção de curvas de crescimento detalhadas que fornecem informações importantes para a identificação de genótipos e direcionamento de programas de melhoramento.

Curvas de crescimento baseadas em curvas logísticas são comumente utilizadas na modelagem de fenômenos biológicos devido à sua eficiência em representar padrões de crescimento. Essas curvas são caracterizadas por um período inicial de crescimento exponencial, que desacelera progressivamente à medida em que se aproxima de um limite máximo (Ornelas et al., 2023). As curvas logísticas podem ser descritas através de seu valor inicial, da inclinação máxima da curva, do ponto no eixo x onde ocorre a inflexão da curva e do deslocamento da curva no eixo y (Zheng et al., 2021).

O ajuste de curvas logísticas a séries temporais de alturas de dosséis fornece uma maneira estruturada de visualizar, entender e comparar os padrões de crescimento e o desempenho de diferentes genótipos e cultivares (Chang et al., 2017). Chang et al. (2017) e Borra-Serrano et al. (2020) usaram curvas logísticas ajustadas a séries temporais geradas via imagens aéreas e SfM para descrever curvas de crescimento e estimar parâmetros relacionados às taxas de crescimento e alturas máximas em lavouras de sorgo e parcelas experimentais de soja, respectivamente.

Os modelos de aprendizado de máquinas permitem que computadores façam previsões com base em correlações e padrões detectados em conjuntos de dados (Ling, 2023), e assim, podem ser aplicados no reconhecimento de padrões nos parâmetros

gerados através da modelagem das curvas de crescimento e classificação de parcelas experimentais quanto ao acamamento.

As aplicações de modelos de aprendizado de máquinas e inteligência computacional na aceleração e maximização dos ganhos de seleção em programas de melhoramento genético de plantas envolvem a análise de grandes conjuntos de dados, reconhecimento de padrões complexos e integração de informações de diversas áreas agronômicas e biológicas (Najafabadi et al., 2023). Diversos tipos de algoritmos de aprendizado de máquina são comumente empregados no melhoramento genético. Os modelos se diferenciam principalmente quanto à necessidade da rotulagem de dados para o treinamento, ao tamanho dos conjuntos de dados necessários para alcançar um treinamento eficaz, na robustez e capacidade de generalização, e na capacidade de executar tarefas classificatórias, de regressão, ou ambas (Govaichelvan et al., 2024; Salokhe, 2023).

Dentre os algoritmos de aprendizado de máquinas disponíveis, os modelos de Random Forest (MRF) se destacam pela consistência, capacidade de generalização e potencial de atingirem altas acurácias mesmo quando treinados com conjuntos de dados relativamente pequenos (Ramezan et al., 2021). Esse método de aprendizado de máquinas supervisionado opera através da agregação das predições de múltiplas árvores de decisão e pode ser utilizado tanto em tarefas classificatórias quanto de regressão. A relativa facilidade de treinamento e capacidade de detectar padrões complexos contribuem para sua ampla adoção na classificação de solos, avaliação da saúde e nutrição de plantas, classificação e agrupamento de espécies vegetais e predição de características agronômicas (Bakhtawer et al., 2022; Devi et al., 2022; Dhillon et al., 2023; Orynbaikyzy et al., 2022).

Sarkar et al. (2023) utilizaram diferentes algoritmos de aprendizado de máquinas para classificar parcelas de soja em níveis de acamamento com base em informações de textura extraída de ortomosaicos gerados através de imagens aéreas. Yang et al. (2021) utilizaram MRF para prever a temperatura do dossel de couve-chinesa a partir de informações meteorológicas e avaliar a condição hídrica da cultura.

O sucesso de MRF depende, dentre outros aspectos, da qualidade e consistência dos conjuntos de dados, da associação das variáveis preditivas com a variável resposta, da representação de classes minoritárias no conjunto de treinamento e do ajuste de hiperparâmetros (Millard & Richardson, 2015; Regier et al., 2023).

Os hiperparâmetros são configurações que determinam as etapas para a execução do treinamento de MRFs e afetam o desempenho do modelo treinado. A otimização dos hiperparâmetros pode resultar em aumento da capacidade preditiva e robustez dos modelos quando o conjunto de dados utilizados no treinamento é pequeno ou desbalanceado e quando as variáveis preditivas apresentam relações complexas. Além disso, o ajuste de hiperparâmetros pode influenciar a capacidade de generalização dos MRFs (Bischl et al., 2023).

De forma análoga às curvas de crescimento, o ajuste de curvas logísticas às séries temporais de alturas de dosséis pode ser utilizado para modelar a ocorrência do acamamento em função da redução da altura. Os parâmetros das curvas ajustadas podem permitir a distinção das parcelas quanto à inclinação das plantas e MRFs podem ser utilizados para identificar esses padrões de forma a classificar parcelas quanto à ocorrência e a severidade do acamamento a partir de dados obtidos via análise de imagens aéreas. A predição do acamamento fundamentada na detecção de reduções nas alturas de dosséis ao longo do ciclo da cultura tem potencial de classificar as parcelas quanto ao acamamento através da detecção indireta do tombamento da haste principal das plantas.

Diante do que foi apresentado, os objetivos gerais do presente trabalho foram avaliar o potencial do uso da análise de imagens aéreas RGB e de modelos de Random Forest gerados via *Surface from motion* na estimação da altura do dossel e da predição do acamamento de parcelas de soja.

Os objetivos específicos foram:

- i) Avaliar a acurácia de diferentes métodos de estimação da altura do dossel a partir de imagens aéreas RGB e SfM no estágio do florescimento.
- ii) Avaliar a acurácia de diferentes métodos de estimação da altura do dossel a partir de imagens aéreas RGB e SfM no estágio da maturação, em ensaios com alta incidência de plantas acamadas.
- iii) Avaliar o potencial de MRF treinados com parâmetros de curvas logísticas ajustadas às séries temporais de alturas de dosséis estimadas a partir de imagens aéreas RGB e SfM em classificar parcelas experimentais de soja quanto à severidade do acamamento.
- iv) Avaliar o potencial de MRF treinados com parâmetros de curvas logísticas ajustadas às séries temporais de alturas de dosséis estimadas a partir de imagens aéreas RGB e SfM em detectar a ocorrência do acamamento em parcelas experimentais de soja.

2. MATERIAL E MÉTODOS

2.1. Material genético

Foi avaliado um painel com 300 linhagens experimentais de soja pertencentes ao banco de germoplasma do Programa de Melhoramento da Soja da Universidade Federal de Viçosa. As linhagens experimentais testadas foram originadas da seleção de plantas individuais na geração F₂, que por sua vez tiveram sua origem em populações obtidas através do cruzamento de cultivares comerciais com características complementares. Os genótipos avaliados apresentavam alta variabilidade de duração do ciclo, com grupos de maturidade relativa estimados entre 6,6 e 8,6, e alta diversidade genética para diversas características de interesse.

As cultivares comerciais GMS 6836 RR, GMS 7069 RR, 7300 NS, 8473 RSF RR, M7739 RR, 8579RSF IPRO e GMS 8365 RR foram incluídas nos ensaios e utilizadas como testemunhas.

2.2. Execução experimental

Dois ensaios compostos pelas mesmas linhagens foram conduzidos na safra de 2020/2021 na Unidade de Ensino, Pesquisa e Extensão Horta Nova, na Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brasil, situada nas coordenadas 20°45'44" S 42°49'23" O, a 663 metros de altitude (Figura 1). Devido à região de cultivo e aos grupos de maturidade relativa, a duração dos ciclos de desenvolvimento dos genótipos variou de intermediária a tardia (Alliprandini et al., 2009). Linhagens de ciclo tardio apresentam maior período vegetativo, o que pode levar ao crescimento excessivo e, consequentemente, maior suscetibilidade ao acamamento (Tomohiro & Homma, 2023).

A semeadura do primeiro ensaio (E1) foi realizada no dia 25 de novembro de 2020, enquanto a do segundo ensaio (E2) foi realizada no dia 07 de dezembro de 2020. Os ensaios foram semeados lado a lado. A diferença de 12 dias entre a semeadura de cada ensaio objetivou proporcionar diferentes condições ambientais durante o desenvolvimento das linhagens de cada ensaio, e assim, permitir a sua caracterização em ambientes diferentes. O delineamento experimental utilizado foi o delineamento de blocos aumentados, com seis blocos. Cada bloco foi composto por 50 linhagens experimentais, avaliadas sem repetições, e por sete testemunhas que se repetiam em cada bloco, o que resultou em um total de 342 parcelas por ensaio.



Figura 1 – Imagem e localização dos ensaios na Unidade de Ensino, Pesquisa e Extensão Horta Nova – UFV. Fonte: Google Earth (2023).

As parcelas de cada ensaio foram dispostas espacialmente no campo em 26 linhas de plantio e 14 colunas. Cada parcela foi constituída por uma linha de 2,5 metros (m), espaçadas em 50 centímetros (cm). Foi realizada a semeadura de 18 sementes m^{-1} , e, após a emergência das plântulas, a densidade de plantas foi ajustada para 12 plantas m^{-1} , o que resultou em uma população de 240.000 plantas por hectare. O manejo dos ensaios foi realizado de acordo com as recomendações para o cultivo da soja na região (Sediyama et al., 2015).

2.3. Avaliações no campo

2.3.1. Características avaliadas

Em cada uma das parcelas, foram avaliadas manualmente, no campo, as seguintes características: número de dias até o florescimento (DIASF, dias), mensurada como o número de dias entre a data do plantio e a data da abertura de pelo menos 50% dos botões florais de 50% das plantas; comprimento de plantas no florescimento (COMPF, cm), mensurado como média da distância entre o ponto de inserção da planta no solo e a extremidade apical da haste principal de duas plantas representativas na data do florescimento da parcela; número de dias até a maturação (DIASM, dias), mensurada como o número de dias entre a data do plantio e a data em que 50% das plantas

apresentavam pelo menos 95% das vagens com coloração característica de vagens maduras; comprimento de plantas na maturação (COMPM, cm), mensurado como média da distância entre o ponto de inserção da planta no solo e a extremidade apical da haste principal de duas plantas representativas na data da maturação da parcela; nível de acamamento (AC), mensurada em uma escala de notas entre 1 e 5, na qual a nota 1 indica plantas predominantemente eretas e a nota 5 indica plantas drasticamente acamadas. A mensuração do acamamento foi realizada com o auxílio de um arco graduado acoplado a uma régua, utilizado para determinar o ângulo de inclinação da haste principal em relação ao solo (Figura 2). Duas plantas representativas foram mensuradas em cada parcela e a média de suas notas foi utilizada para representar a parcela. Em parcelas cujo o cálculo da média resultou em valores decimais, a nota foi arredondada para o valor que indica menor acamamento.

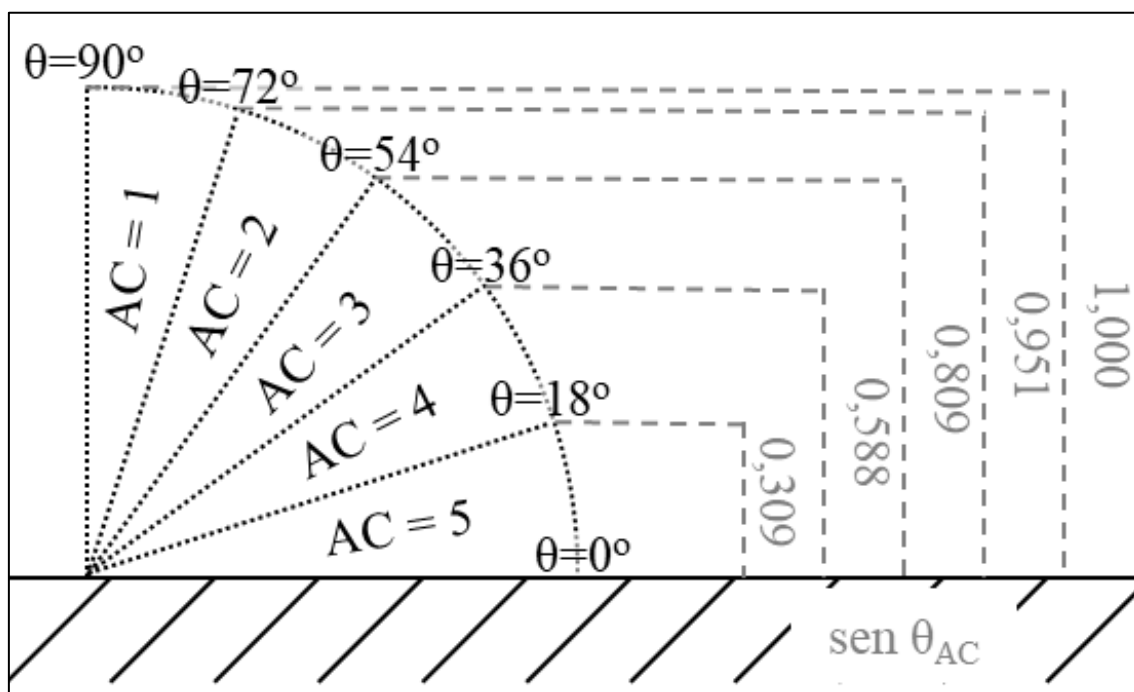


Figura 2 – Diagrama que representa os ângulos de inclinação (θ) que caracterizam os diferentes níveis de acamamento (AC). Os valores de $\text{sen } \theta_{AC}$ indicam a relação esperada entre o comprimento de plantas na maturação e altura do dossel avaliada no campo, em função dos ângulos de inclinação (θ).

A altura do dossel avaliada no campo (ALTDC, cm) foi estimada a partir dos valores de COMPM e AC. A avaliação de AC determinou os ângulos de inclinação máximo e mínimo da haste principal das plantas (Tabela 1). A partir dos ângulos de inclinação, é possível calcular a relação esperada entre o COMPM e ALTDC. Devido a cada AC ser caracterizado por um intervalo contínuo de ângulos de inclinação da haste

principal, foram realizadas duas estimativas de ALTDC para cada parcela, correspondentes aos ângulos mínimo e máximo de cada AC (Tabela 1), de acordo com a equação:

$$\text{ALTDC} = \text{COMPM} \times \sin \theta_{AC}$$

Em que $\sin \theta_{AC}$ representa a relação esperada entre o comprimento de plantas na maturação (COMPM) e a altura do dossel avaliada no campo (ALTDC) de acordo com o ângulo de inclinação da haste principal θ_{AC} .

Tabela 1 – Intervalos de ângulos de inclinação entre a haste principal e o solo (θ) e relação esperada entre o comprimento de plantas na maturação e altura do dossel avaliada no campo de cada nível de acamamento (AC).

AC	Ângulo de inclinação (θ_{AC})	Relação esperada entre COMPM e ALTDC
1	90° a 72°	1,000 a 0,951
2	72° a 54°	0,951 a 0,809
3	54° a 36°	0,809 a 0,588
4	36° a 18°	0,588 a 0,309
5	18° a 0°	0,309 a 0,000

2.3.2. Análise estatística das características avaliadas no campo

O pacote *sommer* (Liaw & Wiener, 2002) do software R (R Core Team, 2023) foi utilizado para fazer o ajuste de modelos lineares mistos às características avaliadas de forma individual para cada ensaio e de forma conjunta.

Nas análises individuais, os modelos completos, com efeitos de genótipos e blocos, considerados aleatórios, foram comparados com os modelos reduzidos, nos quais apenas o efeito dos blocos foi considerado.

Na análise conjunta, foram testadas as significâncias dos genótipos, dos ensaios, e da interação entre os genótipos e os ensaios. O modelo completo compreendeu os efeitos aleatórios dos genótipos, ensaios, blocos, e a interação entre os genótipos e os ensaios. Já os modelos reduzidos foram compostos pelos mesmos efeitos, exceto o efeito a ser testado.

Em todas as análises, a significância dos efeitos foi verificada através do teste da razão da verossimilhança (LRT), com probabilidade igual a 0,05.

2.4. Análise de imagens aéreas

2.4.1. Aquisição de imagens aéreas

A aquisição de imagens aéreas foi realizada com um VANT DJI Phantom 3 Standard acoplado de seu sensor RGB padrão, com 12 megapixels de resolução. A altura de voo foi definida como 18 m acima do solo e a velocidade de voo foi de $2,5 \text{ m s}^{-1}$. As sobreposições frontal e lateral entre imagens foram definidas como 80%. Os voos foram realizados predominantemente em dias ensolarados e sem nuvens, no período entre as 11:00 e às 13:00 horas.

Quatro pontos de controle terrestre (PCT) foram instalados nos vértices da área experimental. Os PCTs foram georreferenciados uma única vez através de um sistema de posicionamento global por pós-processamento e permaneceram fixos na área até a conclusão do último voo. Os PCTs georreferenciados foram utilizados no alinhamento das imagens durante o processamento, e para o alinhamento dos ortomosaicos e dos modelos digitais de elevação (MDE) dos voos sequenciais, realizados ao longo do ciclo da cultura.

Foram executados 19 voos entre 57 e 161 dias após o plantio (DAP) do E1 e 45 e 149 DAP do E2, com intervalo médio de 5,8 dias entre voos (Tabela 2).

Tabela 2 – Lista dos voos executados para a aquisição de imagens aéreas, data de execução e número de dias após o plantio (DAP) em relação a cada um dos ensaios.

Número do Voo	Data de execução	DAP E1	DAP E2
1	21/jan/21	57	45
2	28/jan/21	64	52
3	04/fev/21	71	59
4	11/fev/21	78	66
5	25/fev/21	92	80
6	05/mar/21	100	88
7	18/mar/21	113	101
8	25/mar/21	120	108
9	29/mar/21	124	112
10	01/abr/21	127	115
11	05/abr/21	131	119
12	09/abr/21	135	123
13	16/abr/21	142	130
14	19/abr/21	145	133
15	21/abr/21	147	135
16	27/abr/21	153	141
17	28/abr/21	154	142
18	30/abr/21	156	144
19	05/mai/21	161	149

2.4.2. Processamento de imagens

As imagens de cada data foram processadas através do software Agisoft Metashape para a construção de um ortomosaico (Figura 3a) e de um MDE (Figura 3b).

As imagens foram alinhadas com o auxílio das informações do georreferenciamento dos PCTs. Os ortomosaicos e MDE foram gerados com resoluções espaciais de 0,7 cm pixel⁻¹ e 1,34 cm pixel⁻¹, respectivamente.

O software QGIS foi utilizado em todas as etapas subsequentes do processamento de imagens. A fim de estimar o MDE dos dosséis, foi gerado um MDE do solo de cada MDE original, através da interpolação das regiões que representavam o nível do solo exposto.

Para gerar o MDE do solo, foi calculado o índice de vegetação Excesso de Verde normalizado (ExG) a partir do ortomosaico (Figura 3c), de acordo com a equação proposta por Woebbecke et al. (1995):

$$\text{ExG} = \frac{2 \times G - R - B}{R + G + B}$$

em que G representa o valor da banda da cor verde; R representa o valor da banda vermelho e B representa o valor da banda azul.

Através de múltiplas iterações, foi determinado que valores de ExG inferiores a zero selecionavam satisfatoriamente os pixels que correspondiam ao solo exposto e à palhada, enquanto excluía pixels com vegetação verde. Com base neste critério, uma máscara binária foi gerada com o propósito de desconsiderar as regiões associadas à vegetação verde (Figura 3d).

A máscara gerada a partir do ExG foi utilizada na eliminação dos pixels correspondentes à vegetação verde no MDE (Figura 3e). Apesar disso, foram observadas regiões onde a estimativa da altitude do solo foi superestimada devido à presença de palhada seca sobre o solo, remanescente de uma cultura anterior. Para reduzir a interferência da palhada no MDE do solo, foi realizada uma reamostragem espacial, na qual a resolução espacial foi reduzida para $335 \text{ cm pixel}^{-1}$ (o que representa uma redução de 250 vezes em relação à resolução original do MDE), preservando-se o valor mínimo encontrado nos pixels da área reamostrada (Figura 3f). Caso fossem observadas regiões do MDE do solo com distorções causadas por áreas não descartadas pela máscara ExG mas que não contivessem pixels que representassem o solo, foram aplicados filtro para remover os pixels com inclinação superior a 15%, seguida do interpolação para preencher as lacunas resultantes.

Em seguida, a resolução espacial foi reamostrada para o valor original do MDE através do método de reamostragem por interpolação cúbica B-spline com kernel de tamanho 4x4 (Figura 3g). O raster resultante foi utilizado como MDE do solo e, ao ser subtraído do MDE, foi estimado o MDE dos dosséis (Figura 3h).

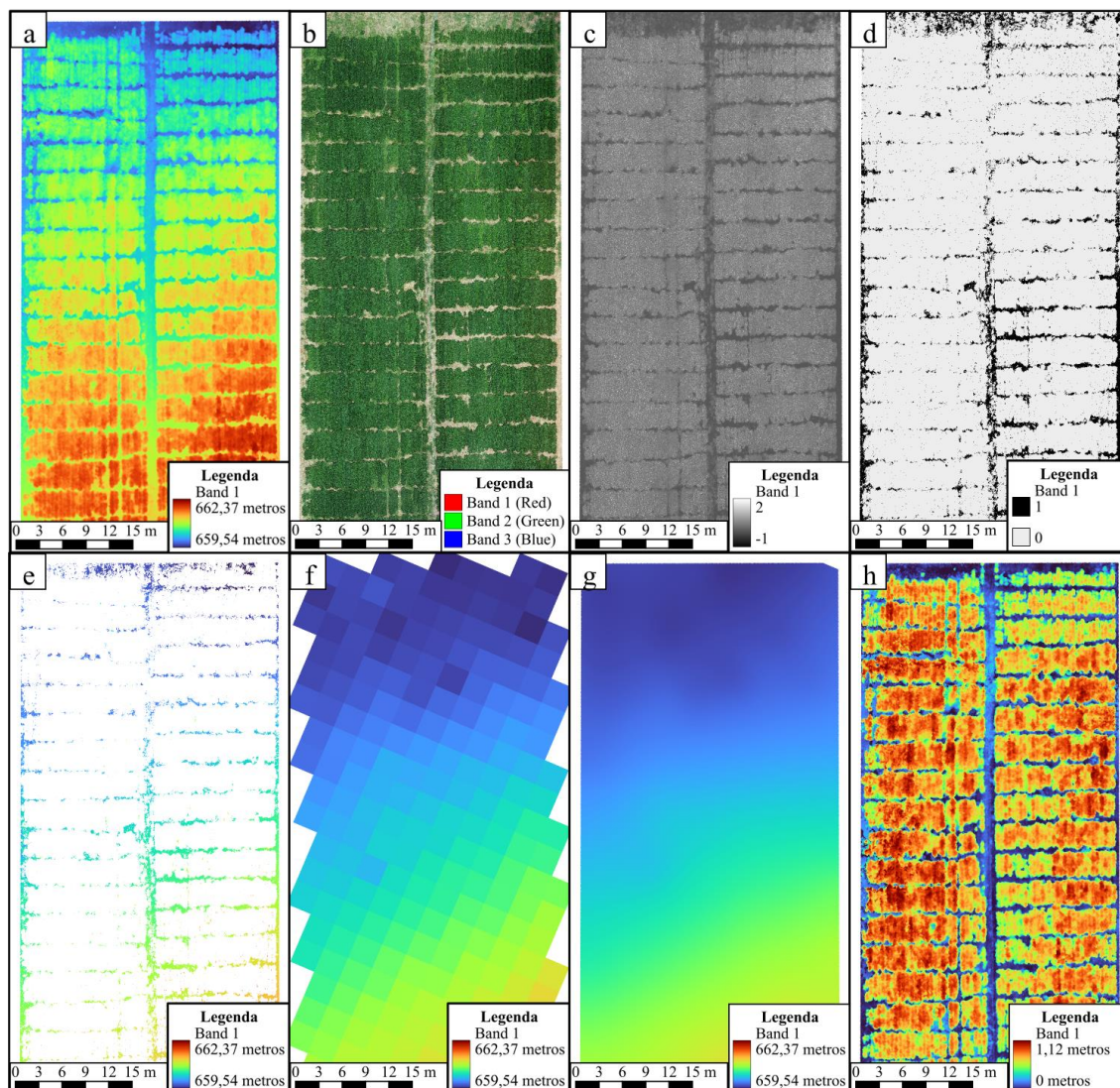


Figura 3 – Etapas do processamento de imagens: (a) Modelo digital de elevação (MDE), gerado a partir do processamento das imagens aéreas; (b) Ortomosaico RGB, gerado a partir do processamento das imagens aéreas; (c) Índice de vegetação Excesso de Verde normalizado (ExG), calculado a partir do ortomosaico RGB; (d) Máscara binária obtida através da seleção dos pixels cujo valor do ExG é igual ou inferior a zero. Pixels com valor igual a um (coloração escura) representam a área sem a presença de vegetação verde; (e) MDE após a eliminação dos pixels associados à vegetação verde; (f) MDE resultante após a redução da resolução espacial e através dos valores mínimos; (g) MDE do solo, gerado através da reamostragem por interpolação cúbica B-spline com kernel de tamanho 4x4; (h) MDE dos dosséis, obtido da diferença entre o MDE e o MDE do solo.

Um vetor de polígonos (*shapefile*) foi posicionado manualmente de forma a sobrepor a região central de cada uma das parcelas. As dimensões de cada polígono foram de 2 m de comprimento e 25 cm de largura (*buffers* de 20% e 50% em relação ao tamanho da parcela, respectivamente). A redução do tamanho dos polígonos em relação às parcelas foi essencial para mitigar a interferência de parcelas vizinhas e dos corredores na

execução das análises, especialmente devido ao potencial tombamento de plantas sobre parcelas adjacentes.

2.5. Estimação da altura do dossel via análise de imagens

As bibliotecas GDAL, NumPy e Pandas, em Python, foram utilizadas para estimar a altura do dossel por análise de imagens (ALTDI, cm) de cada uma das parcelas, a partir dos MDEs dos dosséis de cada voo. A altura do dossel foi estimada nos respectivos percentis da distribuição dos dados, 50% (P50, cm), 75% (P75, cm), 90% (P90, cm), 95% (P95, cm), 99% (P99, cm) e no intervalo de altura predominante (IAP). O IAP foi estimado através da classificação dos pixels em intervalos de dez cm de altura. O número de pixels em cada intervalo foi calculado de acordo com suas alturas, e o intervalo com a maior frequência de pixels foi considerado o valor de IAP. O intervalo de zero a dez cm foi desconsiderado para evitar a interferência da altura do solo e da vegetação espontânea rasteira nas estimativas.

2.5.1. Avaliação das alturas do dossel via análise de imagens

A raiz do erro quadrático médio (RMSE) e a raiz do erro quadrático médio em porcentagem (RMSE%) de cada método de estimação da altura do dossel por análise de imagens (ME_ALTDI) foi utilizada na avaliação da precisão da estimativa do comprimento de plantas no florescimento (COMPF) e altura do dossel avaliada no campo (ALTDC) nos estádios de florescimento e maturação, respectivamente.

O RMSE foi calculado através da seguinte fórmula:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

em que n é o número de observações; y_i são os valores observados e $\hat{y}_i$ são os valores estimados.

O RMSE% foi calculado através da seguinte fórmula:

$$RMSE\% = \left(\frac{RMSE}{y_{\max} - y_{\min}} \right) \times 100$$

em que $y_{\max}$ é a raiz do erro quadrático máxima e $y_{\min}$ é a raiz do erro quadrático mínima.

Devido à não ocorrência de acamamento de nenhuma parcela no período anterior ao florescimento, o valor de COMPF foi utilizado para representar a altura dos dosséis no florescimento. Devido a cada nível de acamamento (AC) ser definido por um intervalo de possíveis ângulos de inclinação da haste principal, foram realizadas duas estimativas

de ALTDC para cada parcela, correspondentes aos ângulos de inclinação mínimo e máximo determinados pela AC da parcela. No estágio da maturação, para cada parcela, foi considerado valor de ALTDC que resultasse na menor raiz do erro quadrático. Nas parcelas em que o valor da altura do dossel estimadas por análise de imagens (ALTDI) se encontrava entre o valor mínimo e máximo da ALTDC, o erro foi considerado zero.

Para minimizar a interferência de variações no valor de COMPF e ALTDC causadas pelo desenvolvimento das plantas, foi considerado o ALTDI estimado a partir do voo com data de execução mais próxima às datas de florescimento e às datas de maturação de cada parcela, respectivamente. Na avaliação de precisão do ALTDI na maturação, foram considerados apenas ALTDI de voos executados anteriormente à data de maturação. Parcelas cuja diferença entre a data do voo mais próximo à data do florescimento ou à data de maturação excedeu sete dias não foram consideradas na avaliação de precisão do ALTDI nos estádios de desenvolvimento do florescimento e da maturação, respectivamente.

A estimativa precisão da altura do dossel estimada por análise de imagens (ALTDI) foram testadas através do RMSE% calculado. A análise de variância foi utilizada para testar a existência de diferenças significativas na distribuição dos RMSE% dos métodos de estimação da altura do dossel por análise de imagens (ME_ALTDI) dentro de cada estágio de desenvolvimento avaliado, entre os níveis de acamamento (AC) dentro de cada ME_ALTDI na maturação, entre os estádios de desenvolvimento dentro de cada ME_ALTDI e entre os melhores ME_ALTDI de cada estágio de desenvolvimento. Os RMSE% dos pares de ME_ALTDI dentro de cada estágio de desenvolvimento avaliado e entre os pares de AC dentro de cada ME_ALTDI na maturação foram avaliados através de testes de comparações múltiplas.

Dois mapas de calor que representam a distribuição espacial das RMSE, estimados através do ME_ALTDI que resultou no menor RMSE%, em cada estágio de desenvolvimento avaliado foram gerados através do pacote R *ggplot2* (Wickham, 2016). O AC foi indicado em cada parcela no mapa de calor do estágio de maturação.

Para possibilitar a comparação entre os estádios de desenvolvimento avaliados, a mesma escala de cores foi aplicada a ambos os mapas de calor. Parcelas com colorações verdes, amarelas e vermelhas representaram parcelas com erros baixos, intermediários e altos, respectivamente. As parcelas omitidas dos mapas de calor foram previamente descartadas ou não dispunham de nenhum voo executado em uma data suficientemente próxima ao DIASF ou DIASM.

2.6. Estimação do acamamento via análise de imagens

2.6.1. Pré-processamento de dados

Para a construção da série temporal de altura do dossel estimadas por meio da análise de imagens (ALTDI), o dia relativo da execução do voo (DIASV) foi calculado a partir do DAP de cada um dos voos e do DIASM de cada uma das parcelas, através da seguinte fórmula:

$$\text{DIASV}_{ij} = \text{DAP}_i - \text{DIASM}_j$$

em que i é o número do voo e j é o número da parcela; DIASV_{ij} é o dia relativo da execução do voo i , na parcela j ; DAP_i é o número de dias após o plantio correspondente à data de execução do voo i e DIASM_j é o número de dias até a maturação da parcela j .

Valores de DIASV superiores a zero indicam que o voo foi executado após a DIASM. As ALTDI estimadas após a maturação não estão associadas ao desenvolvimento das plantas ou ao nível de AC, e dessa forma, não foram incluídos na série temporal.

Em seguida, os valores de ALTDI dos voos com DIASV igual ou inferior a zero foram normalizados pelo valor máximo da série de ALTDI da parcela, através da seguinte fórmula:

$$\text{N_ALTDI}_{ij} = \frac{\text{ALTDI}_{ij}}{\text{ALTDI}_{\text{Max } j}}$$

em que i é o número do voo; j é o número da parcela; N_ALTDI_{ij} é a altura do dossel normalizada pelo valor máximo, estimada através da análise de imagens no voo i , na parcela j ; e $\text{ALTDI}_{\text{Max } j}$ é a ALTDI máxima na parcela j .

Para a execução da normalização pelo valor máximo de IAP, a altura média do intervalo predominante foi adotada como o valor da ALTDI.

A normalização pelo valor máximo torna a ALTDI máxima da parcela, no período considerado, igual a um, e todas as demais em um número entre zero e um proporcional à ALTDI máxima. Esse procedimento evita a interferência direta da altura máxima dos genótipos nas variáveis preditivas dos MRFs treinados posteriormente.

2.6.2. Ajuste da curva logística

Uma curva logística (Yin, 2003) foi ajustada à série temporal de alturas do dossel normalizadas pelo valor máximo (N_ALTDI) de cada parcela através do método de mínimos quadrados não lineares, com utilização do pacote R *minpack.lm* (Elzhov et al., 2016). A equação que descreve a curva ajustada é:

$$N_ALTDI_{ij} = \frac{L_j}{1 + e^{-k_j \times (DIASV_{ij} - x_{0j})}} + b_j$$

em que i é o número do voo; j é o número da parcela; N_ALTDI_{ij} é a altura do dossel normalizada pelo valor máximo estimada através da análise de imagens no voo i , na parcela j ; L_j é o deslocamento máximo da curva ajustada em relação a seu limite inferior; k_j é a inclinação máxima; $DIASV_{ij}$ é o dia relativo da execução do voo i , na parcela j ; x_{0j} é o ponto de inflexão no eixo x e representa o DIASV em que o valor de k_j é máximo; e b_j é o limite inferior da curva ajustada.

A curva de melhor ajuste foi selecionada de acordo com o coeficiente de determinação (R^2) resultante de diferentes combinações dos valores iniciais de L , k , x_0 e b . O número máximo de iterações foi definido como 5.000.

2.6.3. Modelos de Random Forest

Os parâmetros L , k , x_0 , b e R^2 da curva logística ajustada à série temporal de alturas do dossel normalizadas pelo valor máximo (N_ALTDI) de cada parcela foram utilizados como variáveis preditivas para o treinamento de modelos de Random Forest (MRF) classificatórios através do pacote R *randomForest* (Liaw & Wiener, 2002), com o objetivo de prever o nível de acamamento (AC).

2.6.3.1. Otimização de hiperparâmetros

O hiperparâmetro peso de classe permite balancear conjuntos de dados de treinamento com menor frequência de uma ou mais classes. A substituição determina se uma amostra pode ser utilizada pelo modelo de Random Forest (MRF) em várias árvores de decisão. O tamanho dos nós define o critério de parada das divisões dos nós de cada árvore, com base no número de amostras por nó. O número de variáveis aleatórias por divisão controla a diversidade das variáveis preditivas e geralmente têm como valor frações da raiz quadrada do total de variáveis. O número total de árvores especifica quantas árvores serão criadas, de forma que valores maiores aumentam a complexidade do MRF (Moraes & Coelho, 2023; Liaw & Wiener, 2002).

Anteriormente ao início do treinamento, foi realizada a otimização dos hiperparâmetros dos modelos para cada método de estimação da altura do dossel por análise de imagens (ME_ALTDI). Devido ao desbalanceamento das frequências dos AC no conjunto de dados, o hiperparâmetro peso de classes foi definido como o inverso da frequência de cada nível.

Os valores ótimos de substituição, tamanho dos nós, número de variáveis aleatórias por divisão e número total de árvores foram definidos através de uma busca em

grade. A substituição foi testada como TRUE ou FALSE. Para tamanho dos nós, foram testados valores entre um e 20, com incrementos de um. Para número de variáveis aleatórias por divisão, foram testados os valores $\frac{1}{3} \times \sqrt{5}$, $\frac{1}{2} \times \sqrt{5}$, $\frac{3}{4} \times \sqrt{5}$, $\sqrt{5}$, $2 \times \sqrt{5}$ e $3 \times \sqrt{5}$. Para número total de árvores, foi testado o valor um e os valores entre 50 e 2.000, com incrementos de 50. Um total de 9.600 combinações foi testado para cada método de estimação da altura do dossel por análise de imagens (ME_ALTDI).

Após a definição inicial dos hiperparâmetros ótimos, uma nova busca em grade foi realizada para o número total de árvores, na qual os valores adjacentes ao número total de árvores ótimo inicial foram testados com incrementos de um, enquanto os valores dos demais hiperparâmetros foram fixados como os primeiros valores ótimos encontrados. O desempenho de cada combinação de valores foi avaliado através do erro out-of-the-bag. A combinação que resultou no menor erro out-of-the-bag foi selecionada para a utilização no treinamento do MRF.

O erro out-of-the-bag é uma métrica comumente utilizada para a determinação dos valores ótimos de hiperparâmetros de MRF devido à ausência da necessidade da separação de conjuntos de treinamento e de teste (Moraes & Coelho, 2023).

2.6.3.2. Predição do nível de acamamento

Com o objetivo de avaliar sua robustez, dois modelos de Random Forest (MRF) foram treinados para cada ME_ALTDI. O primeiro utilizou o conjunto de dados do ensaio 1 (E1) para treinamento e o do ensaio 2 (E2) para teste, enquanto o segundo utilizou os dados de E2 como conjunto de treinamento e E1 como conjunto de teste. A comparação de MRFs treinados com conjuntos de dados diferentes permite determinar a capacidade de generalização da metodologia e avaliar o potencial de sua aplicação em novos conjuntos de dados. Os pesos de classes foram ajustados de acordo com o inverso da frequência dos AC no conjunto de dados a ser utilizado no treinamento.

Para a visualização da classificação dos conjuntos de teste de cada modelo, foram geradas as matrizes de confusão entre o número de parcelas de cada AC mensurados no campo (AC Real) e seus respectivos números preditos através do modelo de Random Forest (AC Predito) dos dados do conjunto teste. O desempenho geral dos modelos foi avaliado através da acurácia balanceada, coeficiente Kappa e F1-Score. O desempenho por classes foi avaliado através da acurácia balanceada, F1-Score, precisão e sensibilidade. O pacote R *caret* (Kuhn, 2008) foi utilizado para calcular as estatísticas de desempenho dos modelos.

A acurácia balanceada é útil para avaliar modelos treinados com dados desbalanceados, pois considera a média das classificações corretas de cada classe. O coeficiente Kappa mede a concordância entre classificações, considerando o acaso. A precisão reflete a proporção de predições corretas de uma classe, enquanto a sensibilidade mede a capacidade do modelo em identificar todas as instâncias dessa classe. O F1-Score, combinação harmônica de precisão e sensibilidade, avalia o equilíbrio entre falsos negativos e falsos positivos (Florez et al., 2020; Hand et al., 2021). As estatísticas de desempenho apresentadas são úteis para mensurar e comparar a performance dos modelos em identificar corretamente as classes das parcelas avaliadas.

2.6.3.3. Detecção do acamamento via classificação dos níveis de acamamento

Com a finalidade de detectar a ocorrência do acamamento, parcelas com níveis 1 e 2 de AC foram agrupados como “levemente acamados”, enquanto os demais formaram o grupo dos níveis “severamente acamados”. O primeiro grupo foi considerado a classe de interesse na classificação e nas métricas de avaliação do desempenho. As estatísticas de desempenho geral foram utilizadas para avaliar os modelos quanto à detecção do acamamento, de forma análoga à avaliação dos MRFs quanto predição do nível de AC.

A avaliação da capacidade de detecção de acamamento tem o propósito de determinar se os erros de classificação cometidos pelos MRFs originais afetam significativamente as conclusões sobre a resistência ou suscetibilidade dos materiais ao acamamento. Erros classificatórios entre AC associados ao acamamento leve, ou à resistência ao acamamento, geram menor prejuízo para selecionar genótipos resistentes em programas de melhoramento.

O fluxograma das etapas executadas nas avaliações de campo, análise de imagens aéreas, estimação da altura do dossel via análise de imagens e estimação do acamamento é apresentado na Figura 4.

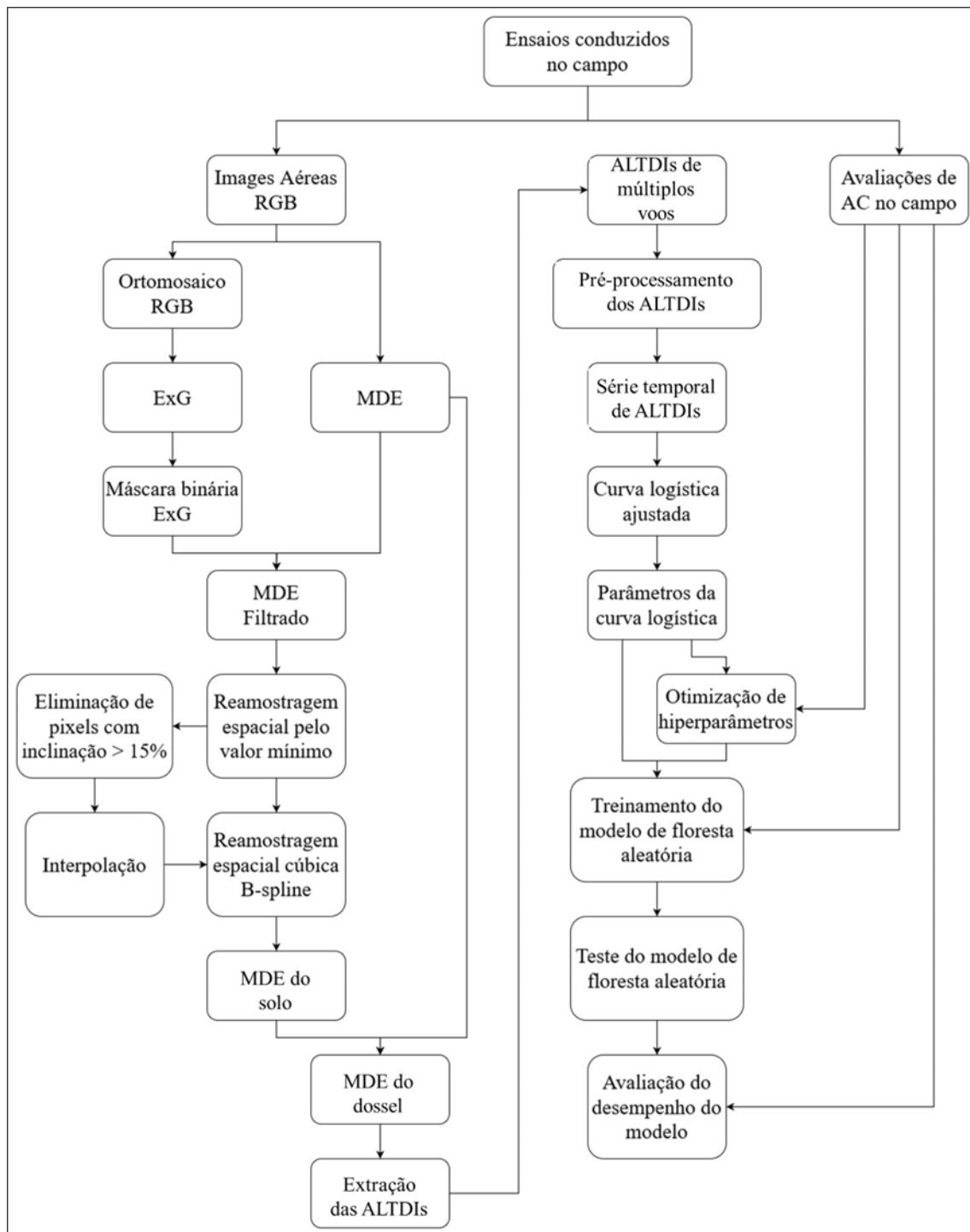


Figura 4 – Fluxograma simplificado das etapas executadas nas avaliações de campo, análise de imagens aéreas, estimação da altura do dossel via análise de imagens e estimação do acamamento.

3. RESULTADOS E DISCUSSÃO

3.1. Avaliações no campo

Do total de 342 parcelas plantadas em cada ensaio, 36 parcelas foram descartadas no E1 e uma no E2. Essas parcelas foram descartadas devido a falhas no plantio e senescência precoce de plantas.

As estatísticas descritivas das avaliações de campo são apresentadas na Figura 5 e na Tabela 3. A altura do dossel avaliada no campo (ALTDC) foi apresentada como a média dos valores calculados a partir dos ângulos de inclinação máximos e mínimos dos níveis de acamamento (AC) correspondentes.

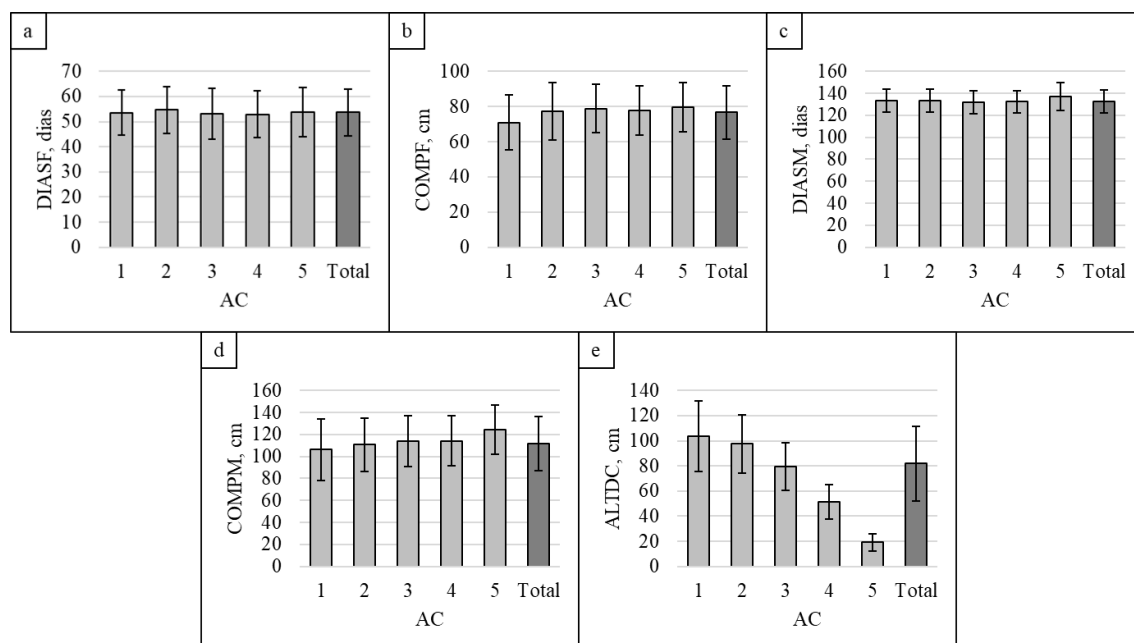


Figura 5 - Média e desvio-padrão das características avaliadas nos ensaios 1 e 2 (E1 e E2, respectivamente) para cada nível de acamamento (AC) e para o conjunto total de dados. As médias são representadas pelas colunas e os desvios-padrão pelos marcadores de dispersão verticais. (a) Número de dias até o florescimento (DIASF); (b) Comprimento de plantas no florescimento (COMPF); (c) Número de dias até a maturação (DIASM); (d) Comprimento de plantas na maturação (COMPM); (e) Altura do dossel avaliada no campo (ALTDC). A ALTDC foi calculada como a média das alturas estimadas a partir dos ângulos de inclinação máximo e mínimo dos respectivos AC.

Tabela 3 – Média e desvio-padrão das características avaliadas nos ensaios 1 e 2 (E1 e E2, respectivamente) para cada nível de acamamento (AC) e para o conjunto total de dados. A altura do dossel avaliada no campo (ALTDC) foi estimada como a média das alturas calculadas a partir dos ângulos de inclinação máximo e mínimo dos respectivos AC.

AC	DIASF – Média (± Desvio- padrão)	COMPF – Média (± Desvio- padrão)	DIASM – Média (± Desvio- padrão)	COMPM – Média (± Desvio- padrão)	ALTDC – Média (± Desvio- padrão)
1	53,5(±8,98)	70,9(±15,53)	133,2(±10,44)	106,1(±27,96)	103,5(±28,00)
2	54,6(±9,17)	77,4(±16,31)	133(±10,41)	110,7(±24,20)	97,5(±23,00)
3	53,1(±10,04)	78,9(±13,78)	131,9(±10,29)	113,8(±22,99)	79,5(±18,67)
4	52,9(±9,29)	77,7(±14,06)	132,3(±9,99)	114,3(±23,07)	51,3(±13,61)
5	53,8(±9,85)	79,6(±14,09)	137,1(±12,96)	124,2(±22,50)	19,2(±7,00)
Total	53,6(±9,41)	76,6(±15,24)	132,7(±10,44)	111,9(±24,64)	81,8(±29,72)

DIASF: número de dias até o florescimento; COMPF: comprimento de plantas no florescimento; DIASM: número de dias até a maturação; COMPM: comprimento de plantas na maturação; ALTDC: altura do dossel avaliada no campo.

A média e o desvio-padrão dos AC foram 2,6 e 1,13, respectivamente. As frequências relativas aos níveis 1, 2, 3, 4 e 5 de AC no conjunto total de dados foram 0,19, 0,29, 0,27, 0,20 e 0,05, respectivamente. A média, o desvio-padrão e as frequências dos AC indicam que a tendência central da característica está levemente deslocada em relação ao nível mediano de AC (nível 3), com uma maior proporção de parcelas classificada nos níveis 2 e 3 de AC, seguida dos níveis 1 e 4 e um menor número de parcelas no nível 5.

As médias de número de dias até o florescimento (DIASF) oscilaram em torno da média geral sem uma tendência clara de variação entre os AC. As médias de comprimento de plantas no florescimento (COMPF) foram similares, exceto no nível 1 de AC, que apresentou menor comprimento de plantas no florescimento.

As médias de número de dias até a maturação (DIASM) também foram similares, exceto no nível 5 de AC, no qual a maturação das parcelas foi mais tardia. As médias de comprimento de plantas na maturação (COMPM) mostram uma tendência de aumento conforme o aumento da severidade dos níveis de AC. Este resultado concorda com os encontrados por Lee et al. (2015), que observaram correlações diretas entre altura de plantas e acamamento, e entre dias até a maturação e acamamento, em avaliações fenotípicas e de associação de QTLs.

Devido aos valores de altura do dossel avaliada no campo (ALTDC) dependerem diretamente dos ângulos de inclinação em relação ao solo que definem AC, suas médias mostraram forte tendência de redução com o aumento da severidade do acamamento.

3.1.1. Análise estatística das características avaliadas no campo

O teste LRT da análise individual dos ensaios mostrou que o efeito dos genótipos foi significativo em todas as características em ambos os ensaios, exceto em DIASF, no E2 (Tabela 4). Apesar disso, a ausência de efeito significativo dos genótipos nessa característica não comprometeu as análises ou resultados posteriores, visto que os valores de DIASF foram utilizados apenas como uma referência temporal para a identificação dos voos realizados em data próxima à data de avaliação do COMPF.

Tabela 4 – Resultados da comparação entre o modelo completo (com efeitos de genótipos e blocos) e o modelo reduzido (sem efeitos de genótipos) no ajuste de modelos lineares mistos para cada característica avaliada em cada ensaio.

Ensaio	Característica	Qui-quadrado	P-valor
E1	DIASF	42,0398	$8,94 \times 10^{-11}$ ***
	COMPF	18,9325	$1,00 \times 10^{-5}$ ***
	DIASM	30,8603	$2,77 \times 10^{-8}$ ***
	COMPM	42,2604	$7,99 \times 10^{-11}$ ***
	AC	8,0791	0,0045 **
	ALTDC	26,8828	$2,16 \times 10^{-7}$ ***
E2	DIASF	1,1110	0,2919 ns
	COMPF	20,1405	$7,20 \times 10^{-6}$ ***
	DIASM	18,8973	$1,00 \times 10^{-5}$ ***
	COMPM	65,9149	$4,71 \times 10^{-16}$ ***
	AC	4,4718	0,0345 *
	ALTDC	9,3752	0,0022 **

E1: Ensaio 1; E2: Ensaio 2; DIASF: número de dias até o florescimento; COMPF: comprimento de plantas no florescimento; DIASM: número de dias até a maturação; COMPM: comprimento de plantas na maturação; AC: nível de acamamento; ALTDC: altura do dossel avaliada no campo; Níveis de significância: ***: P-valor < 0,001; **: P-valor < 0,01; *: P-valor < 0,05; ns: não significativo.

O efeito significativo dos genótipos em AC, ALTDC e COMPF indica que o painel de linhagens selecionado apresenta variabilidade genotípica significativa e é adequado para a avaliação de métodos de estimação de altura do dossel e desenvolvimento de modelos de predição de acamamento. A comparação entre os P-valor do teste qui-quadrado de E1 e E2 sugere que, embora a condução e avaliação dos ensaios tenha sido feita sem nenhuma diferenciação além da data do plantio, as condições ambientais no E1 foram mais consistentes, o que resultou em uma manifestação mais clara das diferenças genotípicas.

O teste LRT da análise conjunta dos genótipos no E1 e E2 (Tabela 5) mostra que o efeito dos genótipos foi significativo para todas as características, exceto DIASF. O efeito do ambiente foi significativo para DIASF, DIASM, AC e ALTDC, e o efeito da

interação entre genótipos e ensaios foi significativo para DIASF, COMPF, DIASM, COMPM e ALTDC.

Tabela 5 – Resultados do teste da razão da verossimilhança (LRT) para os efeitos dos genótipos, ensaios e interação entre genótipos e ensaios das características avaliadas na análise conjunta.

Caract.	Genótipos		Ensaio		Interação entre genótipos e ensaios	
	Qui-Quadrado	P-valor	Qui-Quadrado	P-valor	Qui-Quadrado	P-valor
DIASF	3×10^{-5}	0,9959 ^{ns}	310,90	$1,39 \times 10^{-69}$ ***	12,51	4×10^{-4} ***
COMPF	15,36	$9,00 \times 10^{-5}$ ***	1×10^{-5}	0,9979 ^{ns}	23,19	$1,46 \times 10^{-6}$ ***
DIASM	203,39	$3,80 \times 10^{-46}$ ***	424,51	$2,54 \times 10^{-94}$ ***	5,81	0,0160*
COMPM	282,64	$2,00 \times 10^{-63}$ ***	2,61	0,1065 ^{ns}	13,82	$2,00 \times 10^{-4}$ ***
AC	25,50	$4,43 \times 10^{-7}$ ***	7,39	0,0065**	2,26	0,1328 ^{ns}
ALTDC	45,93	$1,22 \times 10^{-11}$ ***	10,54	0,0012**	7,95	0,0048**

DIASF: número de dias até o florescimento; COMPF: comprimento de plantas no florescimento; DIASM: número de dias até a maturação; COMPM: comprimento de plantas na maturação; AC: nível de acamamento; ALTDC: altura do dossel avaliada no campo; Níveis de significância: ***: P-valor < 0,001; **: P-valor < 0,01; *: P-valor < 0,05; ns: não significativo.

A significância dos efeitos do ambiente e da interação entre genótipos e ensaios indica que os ensaios apresentaram diferenças que influenciaram significativamente os valores médios das características e as respostas dos genótipos em cada ensaio, respectivamente. Isso sugere que os ensaios proporcionaram ambientes distintos para o desenvolvimento das linhagens e que essas diferenças afetaram as características avaliadas.

3.2. Estimação da altura do dossel via análise de imagens

As médias e desvios-padrão das alturas do dossel estimadas por análise de imagens (ALTDI) nos E1 e E2 (Apêndice 1 e Apêndice 2, respectivamente) mostram que, em ambos os ensaios e em todos os métodos de estimação da altura do dossel por análise de imagens (ME_ALTDI), a ALTDI atinge seu valor máximo em torno do voo 5, que corresponde a 92 e 80 DAP para o E1 e E2, respectivamente. Após o voo 5, existe uma tendência geral de redução dos valores de ALTDI até o último voo registrado. A tendência de redução das ALTDI estimadas através do intervalo de altura predominante (IAP) foi mais acentuada em comparação a P50, P75, P90, P95 e P99.

O voo número 5, no qual ocorreu o valor médio máximo de ALTDI, foi executado em um período próximo às DIASF da maioria das parcelas, aos 92 dias após o plantio (DAP) de E1 e 80 DAP de E2 (Tabela 3). Durante a condução dos ensaios, não foi observada a ocorrência do tombamento de plantas anterior ao florescimento em

nenhuma das parcelas. Devido à grande amplitude de DIASM, um número significativo de parcelas atingiu a maturação e permaneceu no campo durante a condução dos ensaios. O acamamento e a maior dificuldade de detecção de plantas após a senescência das folhas e a degradação e quebramento gradual de plantas após a maturação, devido à permanência dos genótipos no campo após sua maturação, podem ter contribuído para a redução da altura observada a partir do voo 5.

As raízes do erro quadrático médio (RMSE) e raiz do erro quadrático médio percentual (RMSE%) foram utilizados para avaliar a concordância entre as ALTDI e as COMPF no estágio de florescimento, e as ALTDI e ALTDC no estágio de maturação. Devido à não-normalidade da distribuição dos resíduos, os testes não-paramétricos de Kruskal-Wallis e de Dunn foram utilizados para testar a significância das diferenças observadas entre os valores de RMSE% dos ME_ALTDI.

Os valores de RMSE e RMSE% no estágio de florescimento (Tabela 6), mostram que o P50 resultou no menor RMSE dentre os ME_ALTDI no florescimento. O teste de Dunn mostrou que os ALTDI estimados como o percentil 50% (P50), IAP e percentil 75% (P75) não se diferem estatisticamente e são superiores aos demais ME_ALTDI quanto à precisão da estimativa do COMPF.

Tabela 6 – Raiz do erro quadrático médio (RMSE) e raiz do erro quadrático médio percentual (RMSE%) dos métodos de estimativa da altura do dossel por análise de imagens (ME_ALTDI) comparados com o comprimento de plantas no florescimento (COMPF), no estágio de florescimento.

	P50	P75	P90	P95	P99	IAP
RMSE	12,9 cm	14,3 cm	16,6 cm	17,9 cm	19,8 cm	13,8 cm
(RMSE%)	(13,5%) a	(15%) a	(17,4%) b	(18,7%) bc	(20,7%) c	(14,5%) a

P50: Percentil 50%; P75: Percentil 75%; P90: Percentil 90%; P95: Percentil 95%; P99: Percentil 99; IAP: Intervalo de altura predominante. RMSE% seguidos pela mesma letra não apresentaram diferenças significativas segundo o teste de Dunn (P-valor > 0,05).

A diferença de altura entre a extremidade apical das hastes principais das plantas e das folhas da porção superior do dossel pode ser parcialmente responsável pelos altos valores de RMSE observados no estágio do florescimento. A altura dessas folhas não foi considerada na estimativa de COMPF, mas suas alturas são representadas nos MDEs dos dosséis. As ME_ALTDI percentil 99% (P99), percentil 95% (P95) e percentil 90% (P90) representam a altura com base nos extremos superiores da altura do dossel, enquanto os valores de P50 e P75 são determinados por porções mais baixas e IAP desconsidera os valores de altura dos extremos superiores das plantas ao ser determinada pela altura da

seção representada com a maior área de superfície vista ortogonalmente. Dessa forma, P50, P75 e IAP tendem a desconsiderar a altura das folhas mais altas e estimam valores de altura do dossel em faixas de altura mais próximas à altura do ponto de crescimento apical das plantas.

Erros na estimação da altura através de modelos digitais de elevação (MDE) gerados por *Surface from motion* (SfM) podem ser causados por fatores inerentes à metodologia e aos cenários reconstruídos. Cenários com pouca textura e características semelhantes e repetitivas podem resultar na correspondência e sobreposição incorreta das imagens, o que impacta a precisão altimétrica da nuvem de pontos em três dimensões e a reconstrução do MDE (Rowell, 2023). Fatores relacionados aos sensores e aos VANTs, como baixa estabilidade do sensor e variações na precisão do georreferenciamento, também podem gerar distorções nas imagens ou nas coordenadas geográficas associadas a elas, o que frequentemente resulta na redução da precisão e confiabilidade dos MDE reconstruídos por SfM (Senn et al., 2022).

Maimaitijiang et al. (2019) utilizaram a combinação de índices de vegetação e MDEs obtidos de imagens aéreas RGB para estimar a altura do dossel e o volume de biomassa em parcelas de soja, e encontraram RMSE% iguais a 11,8% e 14,1%, respectivamente. Han et al. (2018) estimaram a altura de parcelas de sorgo através de MDEs estimados por SfM em cinco diferentes datas. Os autores reportaram RMSEs similares aos do presente trabalho, com variação entre 0,09 e 0,29 m e média 0,19 m. Os autores também relataram que a calibração dos MDE através de uma equação linear calculada a partir de pontos de controle terrestre com altura conhecida gerou reduções de 12,8 a 29,2% nos RMSEs. O uso de fatores de correção é eficiente na amenização de erros sistemáticos nos valores de altura de plantas.

Os valores de RMSE e RMSE%, no estágio de maturação (Tabela 7), mostram que o P99 resultou no menor RMSE dentre os ME_ALTDI na maturação. O teste de Dunn mostrou que, quando o RMSE% do conjunto total das parcelas é comparado independentemente aos níveis AC, os ME_ALTDI P99, P95, P90 e P75 não apresentam diferenças significativas entre si e são superiores a P50 e IAP na maturação.

Tabela 7 – Raiz do erro quadrático médio (RMSE) e raiz do erro quadrático médio percentual (RMSE%) dos métodos de estimação da altura do dossel por análise de imagens (ME_ALTDI) comparados com a altura do dossel avaliada no campo (ALTDC), no estágio de maturação.

AC	RMSE (RMSE%)					
	P50	P75	P90	P95	P99	IAP
1	68,2 cm (57,0%) b	58,8 cm (49,2%) d	52,4 cm (43,8%) c	49,0 cm (41,0%) b	44,5 cm (37,2%) b	72,2 cm (60,4%) b
2	49,8 cm (53,7%) b	39,7 cm (42,8%) b	33,3 cm (35,9%) b	29,8 cm (32,1%) a	25,6 cm (27,6%) a	54,6 cm (58,9%) b
3	33,5 cm (55,9%) b	25,8 cm (43,1%) c	20,8 cm (34,7%) a	19,3 cm (32,2%) a	18,3 cm (30,6%) a	36,6 cm (61,1%) b
4	12,4 cm (20,5%) a	12,5 cm (20,7%) a	16,0 cm (26,4%) a	18,1 cm (29,9%) a	22,5 cm (37,2%) b	16,0 cm (26,4%) a
5	15,1 cm (32,8%) a	20,1 cm (43,7%) c	26,2 cm (57,0%) c	30,3 cm (65,9%) b	37,2 cm (80,9%) c	16,0 cm (34,8%) a
Média	44,4 cm (28,8%) b	36,9 cm (24,0%) a	32,5 cm (21,1%) a	30,6 cm (19,9%) a	28,8 cm (18,7%) a	48,0 cm (31,2%) c

P50: altura do dossel estimada como o percentil 50%; P75: altura do dossel estimada como o percentil 75%; P90: altura do dossel estimada como o percentil 90%; P99: altura do dossel estimada como o percentil 99%; IAP: altura do dossel estimada como o intervalo de altura predominante; RMSE% totais seguidos pela mesma letra não apresentaram diferenças significativas segundo o teste de Dunn (P-valor > 0,05). RMSE% dos níveis de acamamento (AC) dentro dos ME_ALTDI seguidos pela mesma letra não apresentaram diferenças significativas segundo o teste de Dunn (P-valor > 0,05).

O teste H de Kruskal-Wallis entre os RMSE% dos ME_ALTDI no florescimento e na maturação indica que não existem diferenças significativas na precisão das estimativas através do P90 e P95 nos dois estádios de desenvolvimento (P-valor > 0,05). O RMSE% de P50, IAP e P75 foi significativamente menor no estágio de florescimento em relação ao de maturação, enquanto o oposto ocorreu com P99 (P-valor < 0,01).

Também segundo o teste H de Kruskal-Wallis, existem diferenças significativas entre os ME_ALTDI em todos os níveis de AC. O teste de Dunn mostra que métodos que amostram partes inferiores da superfície do dossel, como P50 e IAP, têm maior precisão ao estimar a altura do dossel de parcelas com AC 4 e 5, enquanto métodos que estimam a altura a partir de seções superiores do dossel, como P99 e P95, são mais precisos ao estimar a altura de parcelas AC 2 e 3. Nenhum dos métodos testados se mostrou particularmente eficaz em estimar a altura de parcelas do nível 1 de AC.

A queda de folhas de plantas de soja é uma característica determinante da maturação e está relacionada à remobilização de nutrientes, principalmente do nitrogênio, para o enchimento de grãos (Volpato et al., 2021). Após a senescência e queda das folhas, as parcelas do nível 1 de AC, com plantas predominantemente eretas, são representadas por número de pixels reduzido em comparação às plantas acamadas quando vistas

ortogonalmente, o que torna a resolução espacial um fator crítico para a detecção das plantas não acamadas. Além disso, fatores ambientais, como o vento, podem causar a oscilação das extremidades das plantas capturadas durante os voos de drone. A inconsistência na posição das plantas em imagens adjacentes pode causar erros na reconstrução dos MDE a partir da interpolação das nuvens densas de pontos em três dimensões (Volpato et al., 2021). Obanawa et al. (2020) estimaram a altura de gramíneas através do sensoriamento remoto. Os autores executaram voos em alturas extremamente baixas (2 m) e reportaram irregularidades nos MDE devido ao deslocamento da vegetação causado pelo deslocamento de ar gerado pelas hélices do VANT utilizado.

A subestimação da ALTDI de plantas eretas é coerente com o fato de que P99 mostrou o menor RMSE para as parcelas com nível 1 de AC. Embora muito superiores aos RMSEs calculados no estágio do florescimento, a ALTDI estimada via P99 ameniza o efeito da subestimação da altura do dossel de plantas eretas ao ser definida pelos extremos superiores das alturas dos MDEs dos dosséis, o que resulta na maior aproximação do valor de ALTDC nesse nível.

Borra-Serrano et al. (2020) caracterizaram a cobertura do dossel e a altura do dossel de 454 parcelas de soja através da captura semanal de imagens aéreas ao longo de todo o ciclo da cultura. A cobertura do dossel foi estimada pela área foliar da parcela, detectada através do índice de vegetação ExG, enquanto a altura do dossel foi estimada através do P90. Parâmetros como a altura máxima e taxa de crescimento foram estimados a partir de curvas logísticas ajustadas às séries temporais. Os autores observaram a redução da altura de plantas estimada nos dois voos que antecederam a maturação, executados em um período de 14 dias. A redução da altura estimada foi causada pela senescência das folhas no fim do ciclo de desenvolvimento. Esses resultados concordam com a redução da altura observada no presente trabalho, no qual as ALTDI na maturação foram inferiores às do florescimento, mesmo em parcelas do nível 1 de AC (Apêndice 4).

Plantas de soja severamente acamadas comumente apresentam intensificação do *stay-green*, com retenção de hastes e folhas verdes mesmo após a maturação das vagens (Meyer et al., 2017). Associado ao fototropismo positivo, o *stay-green* faz com que folhas e hastes se posicionem em direção ascendente mesmo durante a maturação, o que ameniza a redução da altura do dossel causada pelo tombamento da haste principal (Morrow et al., 2018). No presente estudo, a retenção de folhas durante e após a maturação foi observada em grande parte das parcelas com acamamento severo, em ambos os ensaios. A detecção da folhagem acima da haste principal pelas imagens aéreas pode ter contribuído

significativamente para a superestimação da ALTDC em parcelas acamadas, o que explica os menores RMSE de P50 em parcelas dos níveis 4 e 5 de AC.

Chu et al. (2018) caracterizaram a altura do dossel de parcelas de milho através de uma série de MDE estimadas via SfM. Os autores relataram que o P99 resultou na melhor estimativa da altura do dossel, medida da base da planta até a extremidade do pendão, com RMSE de 0,21 m no cenário de maior precisão. Os RMSE observados são superiores aos observados no estágio de florescimento e inferiores aos observados no estágio de maturação da soja, no presente trabalho. A não ocorrência do acamamento e características relacionadas à arquitetura e fisiologia de híbridos modernos de milho, como a distribuição das folhas pelo dossel e a retenção das folhas superiores durante a maturação (Huang et al., 2017), podem ter contribuído para a maior consistência e precisão da estimativa da altura via SfM na cultura do milho, em comparação à soja na maturação.

Malambo et al. (2018) estudaram o uso de VANT e técnicas de SfM para estimar a altura de parcelas de milho e sorgo através do P90, P95, P99 e da altura máxima. Eles compararam essa metodologia com a estimação das alturas através de sensores LiDAR em plataformas terrestres e com a mensuração manual no campo. As variações da acurácia foram avaliadas através de voos semanais durante o desenvolvimento das culturas. Os RMSE médios entre alturas mensuradas manualmente e estimadas a partir do SfM, no milho, variaram entre 0,14 e 0,16 m (P99 e altura máxima, respectivamente). Já no sorgo, os RMSE variaram entre 0,18 (P90, P95 e P99) e 0,19 m (altura máxima). O milho mostrou menor estabilidade da acurácia durante o desenvolvimento das plantas, possivelmente devido a diferenças entre as datas de mensurações a campo e da execução dos voos em períodos de rápida elongação das plantas. As estimativas através de SfM e LiDAR apresentaram correlações positivas e significativas, mas as alturas estimadas via SfM foram sistematicamente maiores. As RMSE reportadas indicam acurácia consistente com as observadas no presente estudo no estágio do florescimento, mas foram significativamente superiores quando comparadas às alturas no estágio do acamamento. Em parcelas severamente acamadas, a estimação de ALTDC pode ter sido fonte de erros relacionados à mensuração de AC, visto que a inclinação da base das hastes principais pode não representar perfeitamente a inclinação da planta como um todo.

A comparação entre os ME_ALTDI mais precisos na estimação da altura em cada um dos estádios de desenvolvimento avaliados, P50 no florescimento e P99 na maturação, mostrou que a estimação da altura no florescimento foi mais precisa do que

na maturação ($P\text{-valor} < 0,01$). Esse resultado concorda com a análise dos mapas de calor dos RMSE entre o P50 e COMPF e o P99 e ALTDC, nos estádios de florescimento (Figura 6) e maturação (Figura 7), respectivamente.

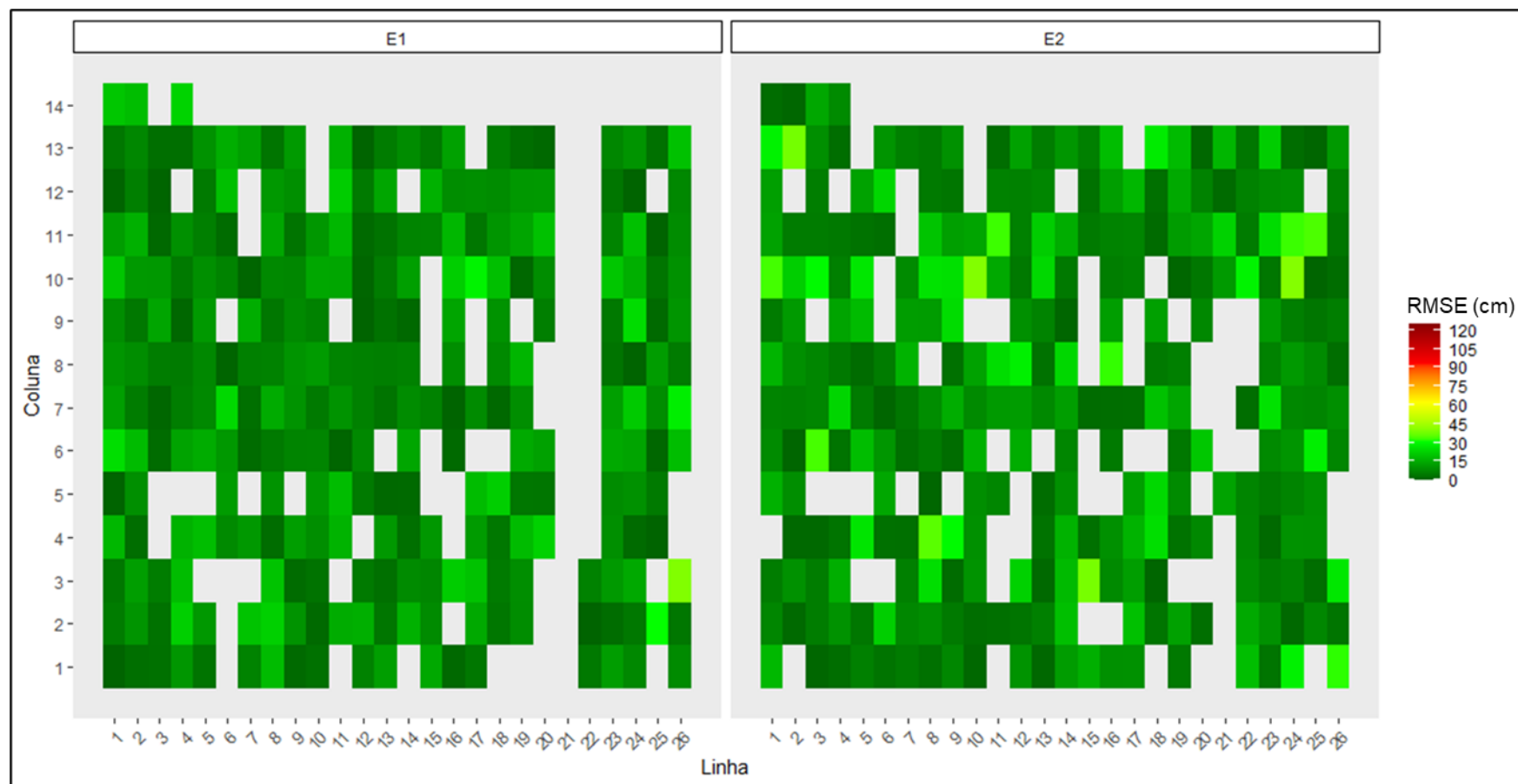


Figura 6 – Mapa de calor da distribuição espacial das raízes do erro quadrático médio (RMSE) no ensaio 1 (E1) e 2 (E2), calculadas a partir dos valores preditos através do percentil 50% (P50) e os comprimentos de plantas no florescimento (COMPF) correspondentes, no estágio de florescimento. As parcelas com coloração verde escuro indicam maior precisão na estimativa de COMPF (menor RMSE) e as parcelas com coloração vermelha indicam menor precisão (maior RMSE).

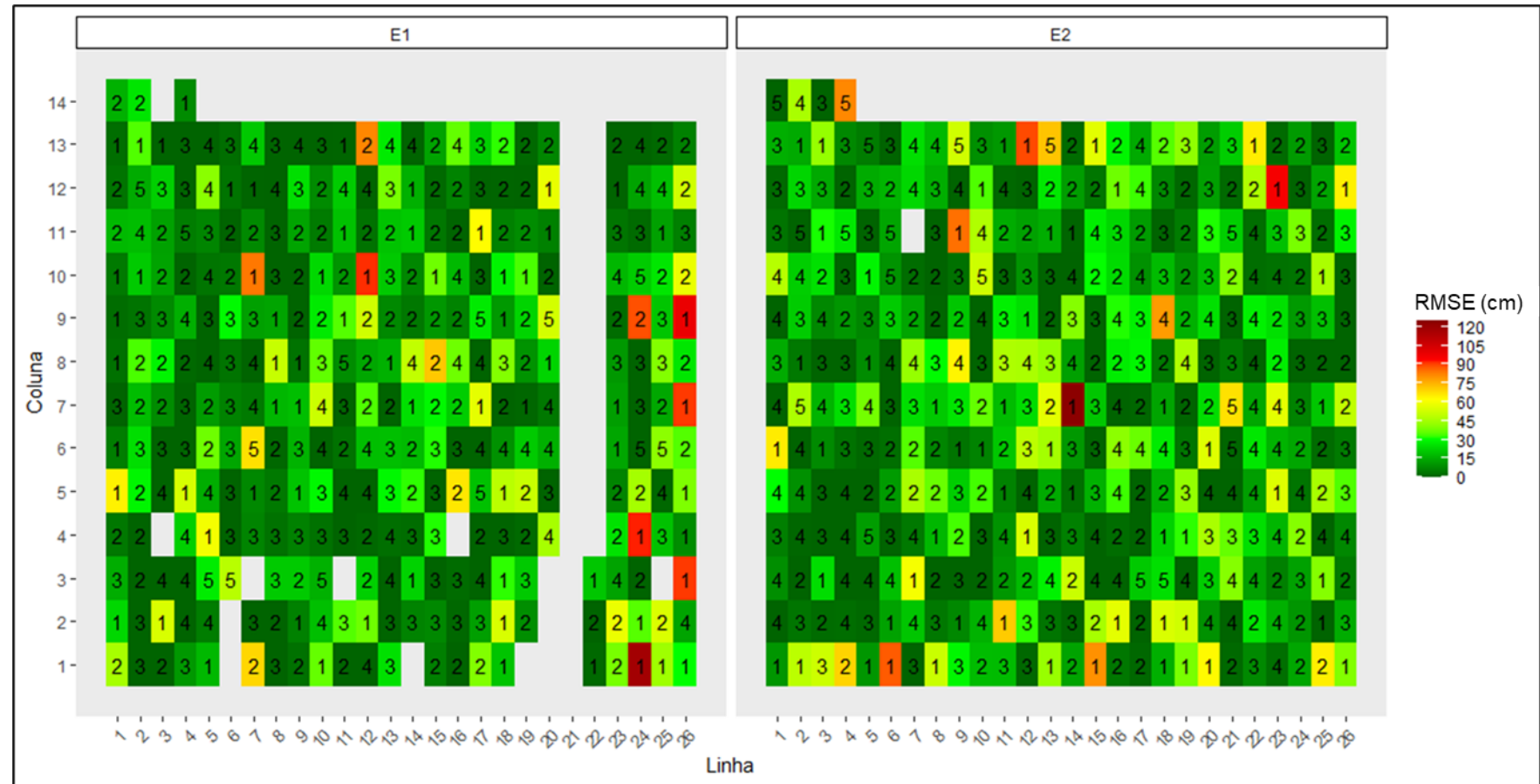


Figura 7 – Mapa de calor da distribuição espacial das raízes do erro quadrático médio (RMSE) no ensaio 1 (E1) e 2 (E2), calculadas a partir dos valores preditos através do percentil 50% (P50) e as altura do dossel avaliada no campo (ALTDC) correspondentes, no estágio de maturação. O número em cada parcela indica seu nível de acamamento (AC). As parcelas com coloração verde escuro indicam maior precisão na estimativa de ALTDC (menor RMSE) e as parcelas com coloração vermelha indicam menor precisão (maior RMSE).

A análise da distribuição dos RMSE entre os níveis de acamamento (AC) mostra que as parcelas com RMSE intermediário, com colorações amareladas, são predominantemente com níveis 4 e 5 de AC, enquanto parcelas com RMSE de alta magnitude, com colorações avermelhadas, são predominantemente parcelas com nível 1 de AC. Os resultados observados nos mapas de calor concordam com os resultados da Tabela 7 quanto à maior precisão da estimação da ALTDC nos níveis 2 e 3 de AC e menor precisão média no nível 1 de AC.

A análise dos mapas de calor de P50 e de P99, no florescimento e na maturação, respectivamente, revela que a intensidade e frequência dos RMSE não apresenta padrões claros de tendência espacial na área experimental. A distribuição dos RMSE é consistente entre os ensaios e não mostra agrupamentos significativos de alta ou baixa precisão. Isso evidencia que a metodologia adotada para a estimação dos MDEs do solo não sofreu fortes distorções espaciais e foi eficaz em não introduzir vieses espaciais nos ALTDI estimados a partir dos MDEs dos dosséis.

3.3. Estimação do acamamento via análise de imagens

3.3.1. Série temporal de alturas do dossel normalizadas pelo valor máximo

As séries temporais de alturas do dossel normalizadas pelo valor máximo (N_ALTDI) em função do dia relativo da execução do voo ($DIASV$) (Figura 8) indicam a existência de comportamentos médios distintos entre os níveis de AC. Em todos os métodos de estimação da altura do dossel por análise de imagens (ME_ALTDI), a série temporal da média das N_ALTDI de cada nível de acamamento (AC) demonstrou um período de crescimento até cerca de -63 dias, seguido por um período de estabilidade e de declínio. No nível 1 de AC, o período de estabilidade durou de -63 até -28 dias, seguido por um forte declínio até a maturação. Nos demais níveis, o período de estabilidade foi mais curto e praticamente inexistente nos níveis 2 e 5 de AC. Nota-se também que a velocidade do declínio das alturas do dossel estimadas por análise de imagens ($ALTDI$) foi menor à medida que o acamamento é mais severo. O declínio mais brando e com maior duração nos níveis de acamamento mais severo pode estar relacionado ao efeito de redução da média da N_ALTDI ao longo de todo o período estudado causado pelo acamamento das parcelas em diferentes $DIASV$.

A N_ALTDI média na maturação mostra uma tendência de alturas levemente maiores à medida que o AC diminui, com exceção do nível 1 de AC. A N_ALTDI média

na maturação do nível 1 de AC não seguiu a tendência geral e foi equivalente à do nível 4 de AC.

A redução da área da superfície que determina a altura das plantas causada pela queda das folhas apresenta desafios adicionais para a estimação precisa das ALTDI. A resolução espacial é crucial para a detecção de objetos pequenos em cenários complexos (Casella et al., 2022; Kalimuthu et al., 2016). Quando utilizada para estimar a altura de objetos com pequenas áreas em relação à altura, fatores como a falta de retenção de detalhes do objeto causada por limitações da resolução espacial, variações altimétricas abruptas no cenário e baixa distinção entre o objeto de interesse e o cenário ao seu redor podem levar à suavização da nuvem de pontos em três dimensões e do modelo digital de elevação (MDE), e consequentemente, à subestimação altimétrica (Jian et al., 2022; Li et al., 2021).

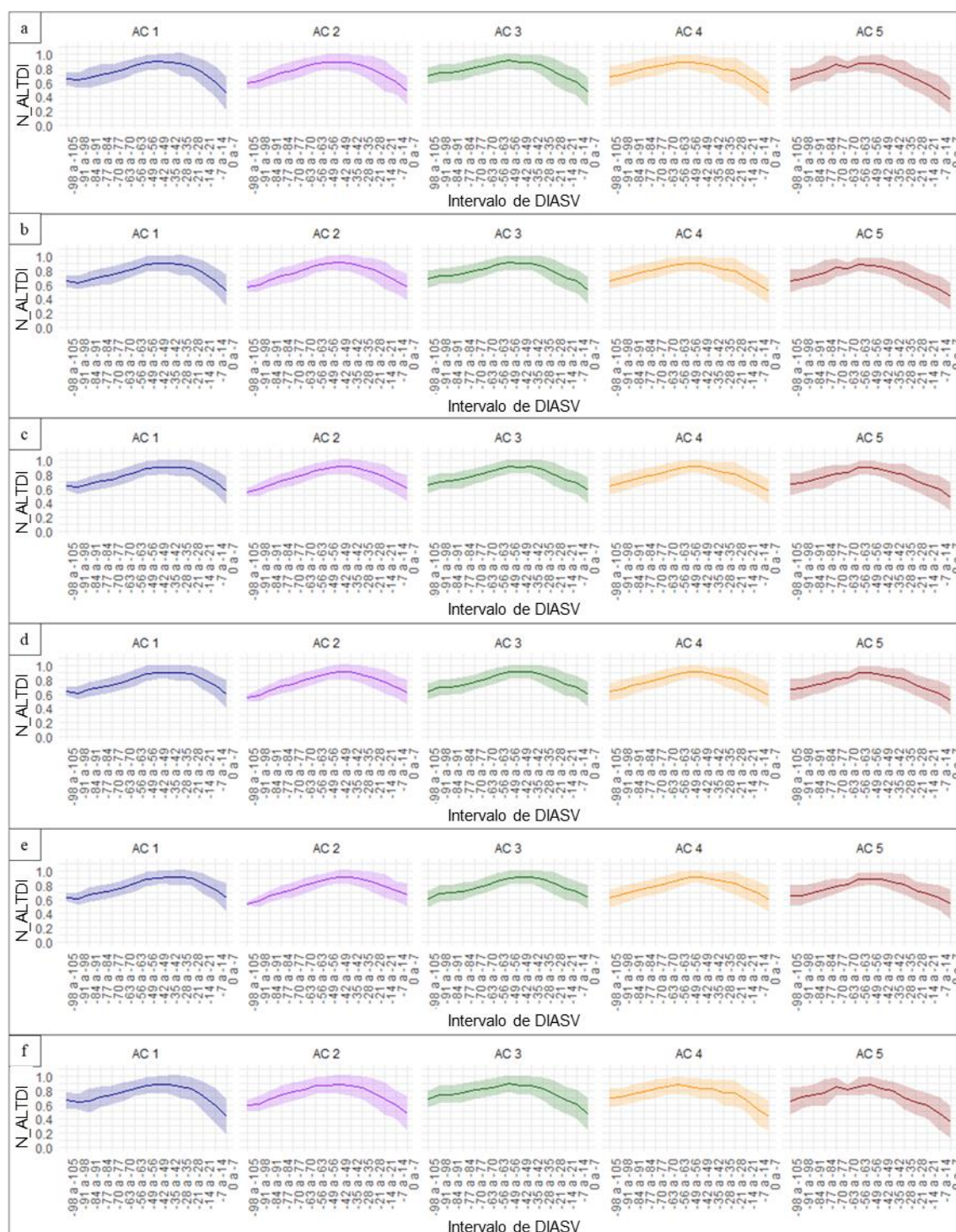


Figura 8 – Séries temporais de alturas do dossel normalizadas pelo valor máximo (N_ALTDI) em intervalos de 7 dias relativo da execução do voo (DIASV), entre os DIASV -105 e 0. Os níveis de acamamento (AC) são apresentados horizontalmente em colunas. As linhas sólidas representam a média de N_ALTDI no intervalo de DIASV, enquanto as áreas sombreadas indicam o desvio-padrão. (a) Série temporal do percentil 50% (P50). (b) Série temporal do percentil 75% (P75). (c) Série temporal do percentil 90% (P90). (d) Série temporal do percentil 95% (P95). (e) Série temporal do percentil 99% (P99). (f) Série temporal do intervalo de altura predominante (IAP).

Apesar da alta resolução dos ortomosaicos e dos MDE, teoricamente suficientes para a detecção dos racemos apicais de vagens maduras, a detecção precisa da altura do dossel em plantas eretas ou levemente acamadas após a senescência e queda das folhas pode ter sido limitada pelo conjunto de fatores relacionados ao sensor, à metodologia de reconstrução dos MDE, à complexidade dos cenários e às influências ambientais.

3.3.2. Ajuste da curva logística

O modelo de ajuste da curva logística convergiu com sucesso para todas as 647 parcelas de E1 e E2. A Figura 9 mostra a curva ajustada às médias das N_ALTDIs de cada DIASV, nos diferentes níveis de AC. Embora exista certa sobreposição das curvas ajustadas às médias de N_ALTDI nos níveis 2, 3 e 4 de AC em todos os métodos de estimação da altura do dossel por análise de imagens (ME_ALTDI), a análise visual evidencia um comportamento médio distinto entre os AC em todos os ME_ALTDI.

De maneira geral, os AC associados ao acamamento leve apresentam inclinação máxima mais acentuada (maiores valores absolutos de k) e o ponto de inflexão da curva é mais próximo ao DIASM (valor de x_0 é mais próximo de 0) quando comparados a AC que representam acamamento severo. As diferenças entre valores médios de k e x_0 são particularmente evidentes no intervalo de altura predominante (IAP) e percentil 95% (P95), respectivamente.

A análise dos intervalos definidos pelos primeiro e terceiro quartis (Apêndice 3) evidencia grande sobreposição de valores de todos os parâmetros estimados entre os AC, independentemente do ME_ALTDI. Os valores distintos das medianas e a sobreposição dos valores do primeiro e terceiro quartil entre os AC indicam que, embora existam padrões de comportamento entre as medianas dos parâmetros e a relação esperada entre o comprimento de plantas na maturação e altura do dossel avaliada no campo, a grande dispersão dos valores dos parâmetros dentro dos AC impedem que uma associação similar seja estabelecida quando são consideradas as curvas ajustadas às séries temporais de cada parcela, individualmente.

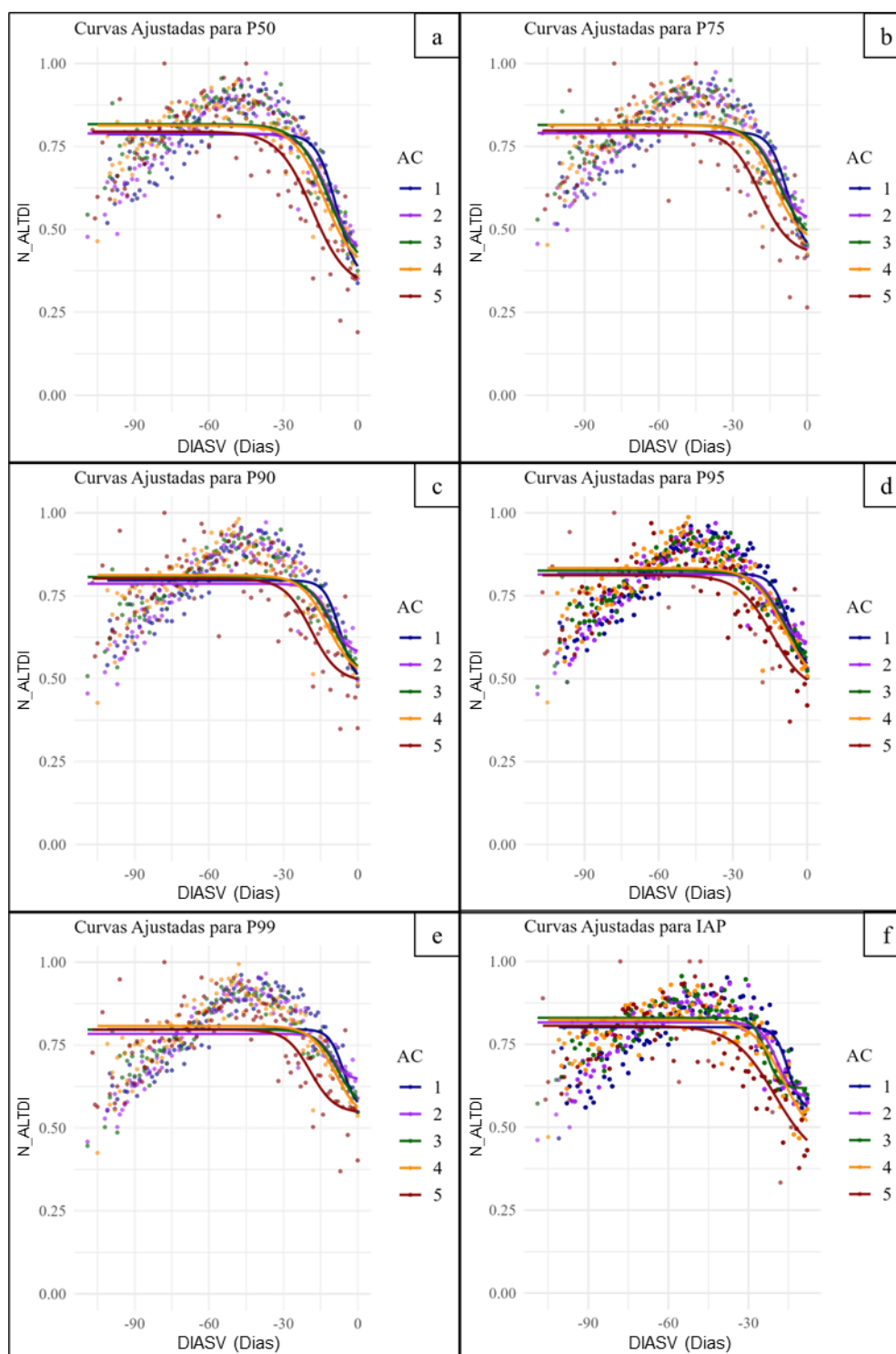


Figura 9 – Curvas logísticas ajustadas à série temporal dos níveis de acamamento (AC), a partir das médias de altura do dossel normalizadas pelo valor máximo (N_ALTDI) em cada dia relativo da execução do voo ($DIASV$). O valor 0 no eixo $DIASV$ representa a data da maturação das parcelas. (a) Curvas ajustadas ao percentil 50% (P50). (b) Curvas ajustadas ao percentil 75% (P75). (c) Curvas ajustadas ao percentil 90% (P90). (d) Curvas ajustadas ao percentil 95% (P95). (e) Curvas ajustadas ao percentil 99% (P99). (f) Curvas ajustadas ao intervalo de altura predominante (IAP).

Embora não tenham sido encontrados artigos buscam prever o acamamento através do ajuste de curvas logísticas a séries temporais de alturas estimadas através de análise de imagens aéreas, metodologias semelhantes foram utilizadas para estimar curvas de crescimento de culturas agrícolas.

Chang et al. (2017) realizaram sete voos durante o ciclo de desenvolvimento de uma lavoura de sorgo. As imagens RGB capturadas foram utilizadas para gerar uma série de MDE através de *Surface from motion* (SfM). O modelo digital do dossel foi estimado através da subtração de um MDE do solo dos MDEs originais. A área foi dividida em uma grade e a altura máxima de cada célula foi utilizada para calcular a altura média, considerada a altura do dossel. A série temporal de alturas do dossel foi utilizada para o ajuste de uma curva sigmoide. A curva ajustada às alturas do dossel estimadas via análise de imagem superestimou as alturas reais, mensuradas no campo, principalmente durante o período de maior taxa de crescimento da cultura. A raiz do erro quadrático médio (RMSE) entre a altura estimada via análise de imagens e as alturas reais foi de 0,33 m.

Diferentemente do presente trabalho, no qual a estimação das ALTDI se iniciou no período de florescimento e as curvas logísticas ajustadas foram majoritariamente decrescentes, Borra-Serrano et al. (2020) construíram a série temporal das alturas do dossel estimadas a partir de imagens adquiridas semanalmente em um ensaio com parcelas de soja, ao longo de todo o ciclo da cultura, e fizeram o ajuste de uma curva logística crescente aos dados. Embora não tenha sido reportada a ocorrência de acamamento, os autores observaram a redução da altura do dossel estimada nos dois voos que antecederam a maturação, causada pela senescência natural das plantas no fim de seu ciclo. As informações obtidas nos últimos dois voos foram desconsideradas para evitar desvios na representação do crescimento nas curvas ajustadas. A desconsideração das informações dos voos realizados nos 14 dias para evitar a subestimação das ALTDI no período que antecede a maturação, não é adequada aos objetivos do presente trabalho, visto que, desconsiderar essas informações também desconsideraria o acamamento que ocorre nos dias antecedentes à maturação.

As interferências reportadas por Borra-Serrano et al. (2020) também foram observadas no presente trabalho. O decréscimo acentuado das ALTDI das parcelas com níveis 1 e 2 de AC se dá a partir dos DIASV -15 e -30, respectivamente. A interferência da maturação nas ALTDI pode ser responsável por grande parte da sobreposição dos parâmetros das curvas ajustadas. A subestimação das ALTDI de parcelas levemente acamadas devido à baixa precisão na detecção das extremidades das hastes principais

após a senescência e queda das folhas devido à maturação leva ao ajuste de curvas semelhantes às de parcelas cuja redução da ALTDI no período próximo à maturação foi decorrente do acamamento.

3.3.3. Modelos de Random Forest

Os modelos de Random Forest (MRF) foram treinados com os parâmetros do ajuste da curva logística à série temporal de N_ALTDIs como variáveis preditoras, e AC como a variável resposta. Dois conjuntos de dados foram utilizados: no primeiro modelo, os dados de ensaio 1 (E1) foram utilizados para o treinamento e os do ensaio 2 (E2) para teste e, no segundo modelo, os dados de E2 foram utilizados para o treinamento e os de E1 para teste. Devido às diferenças no comportamento dos genótipos quando avaliados no E1 e no E2 (Tabela 5), a utilização desses conjuntos de dados separadamente no treinamento e teste permitem avaliar a robustez da metodologia quando aplicada a conjuntos de dados distintos.

O procedimento foi realizado individualmente para os métodos de estimação da altura do dossel por análise de imagens (ME_ALTDI) P50 (MRF_P50), P75 (MRF_P75), P90 (MRF_P90), P95 (MRF_P95), P99 (MRF_P99) e IAP (MRF_IAP).

3.3.3.1. Otimização de hiperparâmetros

Os valores de pesos de classes (Tabela 8) foram definidos como o inverso da frequência da ocorrência dos AC no conjunto de dados total e, dessa forma, foram mantidos fixos durante a otimização de hiperparâmetros em todos os MRFs.

Tabela 8 – Níveis de acamamento (AC), frequência da ocorrência no conjunto de dados e valores de pesos de classes adotados durante o treinamento dos modelos de Random Forest (MRF) para a otimização de hiperparâmetros.

AC	Frequência no conjunto de dados	Peso de classe
1	0,19	5,13
2	0,29	3,44
3	0,27	3,70
4	0,20	5,05
5	0,05	21,57

O ajuste do hiperparâmetro pesos de classes é essencial para a robustez dos MRFs quanto ao desbalanceamento de classes no conjunto de treinamento. A Tabela 9 apresenta os valores de substituição, tamanho dos nós, número de variáveis aleatórias por divisão e número total de árvores que resultaram no menor erro out-of-the-bag durante a

otimização de hiperparâmetros. Os valores ótimos selecionados foram utilizados no treinamento do MRF.

Tabela 9 – Valores ótimos dos hiperparâmetros substituição, tamanho dos nós, número de variáveis aleatórias por divisão e número total de árvores para cada modelo de Random Forest (MRF) determinados através da busca em grade.

Modelo	Substituição	Tamanho dos nós	Número de variáveis aleatórias por divisão	Número total de árvores	Erro out-of-the-bag
MRF_P50	TRUE	7	$\frac{1}{2} \times \sqrt{5}$	50	0,6422
MRF_P75	FALSE	5	$3 \times \sqrt{5}$	60	0,6158
MRF_P90	FALSE	1	$\sqrt{5}$	21	0,6452
MRF_P95	TRUE	2	$\sqrt{5}$	97	0,6276
MRF_P99	TRUE	3	$2 \times \sqrt{5}$	65	0,6716
MRF_IAP	TRUE	1	$2 \times \sqrt{5}$	43	0,6745

MRF_P50: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 50%; MRF_P75: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 75%; MRF_P90: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 90%; MRF_P95: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 95%; MRF_P99: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 99%; MRF_IAP: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o intervalo de altura predominante.

3.3.3.2. Predição do nível de acamamento

Os pesos de classes utilizados nos treinamentos com o conjunto de dados de E1 e E2 são apresentados na Tabela 10.

Tabela 10 – Níveis de acamamento (AC), frequência da ocorrência no conjunto de dados e valores de pesos de classes adotados no treinamento dos modelos de Random Forest (MRF) treinados com os conjuntos de dados do ensaio 1 (E1) e 2 (E2).

AC	Treinamento com E1 e teste com E2		Treinamento com E2 e teste com E1	
	Frequência no conjunto de dados de treinamento	Peso de classe	Frequência no conjunto de dados de treinamento	Peso de classe
1	0,22	4,50	0,17	5,88
2	0,33	3,06	0,26	3,88
3	0,25	4,08	0,29	3,41
4	0,16	6,12	0,23	4,37
5	0,04	23,54	0,05	20,06

O desbalanceamento nas frequências de classes nos conjuntos de dados de treinamento pode ter impacto negativo no desempenho do MRF ao causar o favorecimento de classes majoritárias e o desprezo das classes minoritárias (Hairani et al., 2023; Johnson & Khoshgoftaar, 2019). A atribuição de pesos às classes em conjuntos

de dados desbalanceados garante que o modelo busque determinar os padrões das classes minoritárias de forma equivalente às classes majoritárias, o que melhora o desempenho da classificação (Bakirarar & Elhan, 2023).

As matrizes de confusão (Tabela 11) mostram grande superestimação do número de parcelas pertencentes ao nível 1 de acamamento (AC) nos modelos treinados a partir das alturas do dossel estimadas por análise de imagens (ALTDI) como os percentis 50% (MRF_P50) e 75% (MRF_P75), e em todos os MRFs treinados com o conjunto de dados do ensaio 2 (E2). Nos MRF_P50 e MRF_P75 treinados com o conjunto de dados do ensaio 1 (E1), houve ainda a subestimação do número de parcelas com níveis 2 e 3 de AC devido à sua classificação incorreta no nível 5. Os MRF_P90 a MRF_P99, quando treinados com o conjunto de dados de E1, subestimaram o número de parcelas com níveis 1 e 5 ao classificá-las nos níveis 3 e 4.

A ausência de padrões claros e a sobreposição das variáveis preditoras de diferentes classes podem levar os modelos a identificar, no treinamento, padrões incorretos ou irrelevantes, o que introduz ruído e viés na classificação do conjunto de teste (Malhotra & Karanickolas, 2020). Embora existam distinções no comportamento médio das curvas ajustadas às parcelas de cada AC (Figura 9), a sobreposição dos parâmetros da curva ajustada de parcelas de diferentes AC (Apêndice 3) pode ter impossibilitado a classificação correta das parcelas com parâmetros que não correspondem ao padrão médio de sua classe.

Quando treinados com o conjunto de dados do ensaio 1, os MRFs mostraram tendência de subestimação nos níveis 1 e 2 de AC, e quando treinados com o conjunto de dados de E2, os MRFs superestimaram o nível 1 e subestimaram os níveis 3 e 5 de AC. As matrizes de confusão mostram que, em geral, houve grande confundimento entre os níveis 1, 2, 3 e 4 de AC. Embora os MRFs tenham mostrado seu pior desempenho na classificação das parcelas no nível 5 de AC, a classe não foi ignorada. O número de falsos positivos e negativos indica que a baixa capacidade preditiva neste nível está associada à grande amplitude dos valores de parâmetros chave para a sua classificação, o que não permitiu a identificação de um padrão característicos dessas parcelas.

A média das estatísticas de desempenho geral mostra que, embora o conjunto de dados de E1 ser maior, os modelos treinados com E1 e E2 apresentaram um desempenho médio similar. Isso indica que, embora não tenha sido determinado um padrão claro e distinto entre os níveis de AC, o comportamento dos parâmetros das curvas ajustadas é consistente entre os ensaios. Em geral, as acurácias balanceadas e os índices Kappa dos modelos treinados com conjuntos de dados de E1 e E2, que indicam as frequências de classificações corretas dos modelos e a concordância das classificações, respectivamente, concordaram entre si quanto ao desempenho das classificações de todos os ME_ALTDI. Embora os F1-Score médios com o treinamento feito a partir de E1 e E2 tenham sido similares, as métricas dos MRF_P50 apresentaram variações com a inversão do conjunto de treinamento (Tabela 12).

Tabela 12 – Estatísticas de desempenho geral acurácia balanceada, F1-Score e coeficiente Kappa dos modelos de Random Forest (MRF) treinados com os conjuntos de dados obtidos por diferentes métodos de estimação da altura do dossel por análise de imagens dos ensaios 1 (E1) e 2 (E2).

MRF	Treinamento com E1 e teste com E2			Treinamento com E2 e teste com E1		
	Acurácia Balanceada	F1-Score	Coeficiente Kappa	Acurácia Balanceada	F1-Score	Coeficiente Kappa
MRF_P50	0,5611	0,2647	0,1031	0,5376	0,2072	0,0721
MRF_P75	0,5140	0,1990	0,0159	0,5207	0,2196	0,0434
MRF_P90	0,5184	0,2239	0,0421	0,5285	0,2210	0,0472
MRF_P95	0,5371	0,2481	0,0543	0,5114	0,2203	0,0050
MRF_P99	0,5200	0,2377	0,0281	0,5225	0,2300	0,0493
MRF_IAP	0,4996	0,1983	0,0000	0,4987	0,1899	0,0000
Média	0,5250	0,2286	0,0406	0,5199	0,2147	0,0362

MRF_P50: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 50%; MRF_P75: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 75%; MRF_P90: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 90%; MRF_P95: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 95%; MRF_P99: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 99%; MRF_IAP: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o intervalo de altura predominante.

O F1-Score combina a precisão e a sensibilidade em uma única métrica, e assim indica o equilíbrio de modelos classificatórios quanto a minimizar falsos negativos e falsos positivos (Hand et al., 2021). O treinamento com E1 resultou em estimativas superiores para ambas as métricas que compõe o F1-Score (Apêndice 5), o que indica que, quando comparado a E2, as curvas logísticas ajustadas ao conjunto de dados de E1 forneceram padrões que favoreceram a distinção entre as parcelas no nível 1 de AC e as dos demais níveis. A identificação desses padrões pelo MRF levou à redução dos falsos negativos e positivos na predição das parcelas dessa classe.

Os modelos treinados a partir das ALTDI como o intervalo de altura predominante (MRF_IAP) apresentaram os piores desempenhos entre os métodos avaliados. O coeficiente Kappa indica que o modelo não teve poder preditivo adequado e falhou em identificar quaisquer padrões distintivos entre os níveis de AC.

Em média, não houve grandes variações entre as acurácias balanceadas nos diferentes AC (Apêndice 5). O nível 1 e o nível 3 resultaram na melhor e na pior acurácia balanceada média, respectivamente. O nível 5 apresentou F1-Score, precisão e sensibilidade consideravelmente inferiores aos dos demais AC, o que indica que os modelos tiveram baixa eficiência em evitar falsos negativos e falsos positivos neste nível.

Consequentemente, os modelos não conseguiram identificar padrões que permitissem a distinção do nível 5 em relação aos outros níveis.

As acurácias balanceadas dos AC do MRF_P50, cujo desempenho foi superior aos demais MRF, indicam haver diferenças na acurácia classificatória nos diferentes AC (Figura 10). O nível 1 apresentou a melhor acurácia balanceada, o maior F1-Score, e a maior sensibilidade. A sensibilidade deste nível indica que os MRF_P50 foram capazes de identificar e classificar a maioria das parcelas deste nível na classe correta. No entanto, o modelo não se destacou quanto à precisão, o que indica que um número significativo de parcelas de outros níveis foi incorretamente classificado como pertencente ao nível 1, o que concorda com a superestimação do número de parcelas na classe (Tabela 11). O nível 4 de AC também se destacou positivamente em relação aos demais níveis, especialmente em sensibilidade e F1-Score.

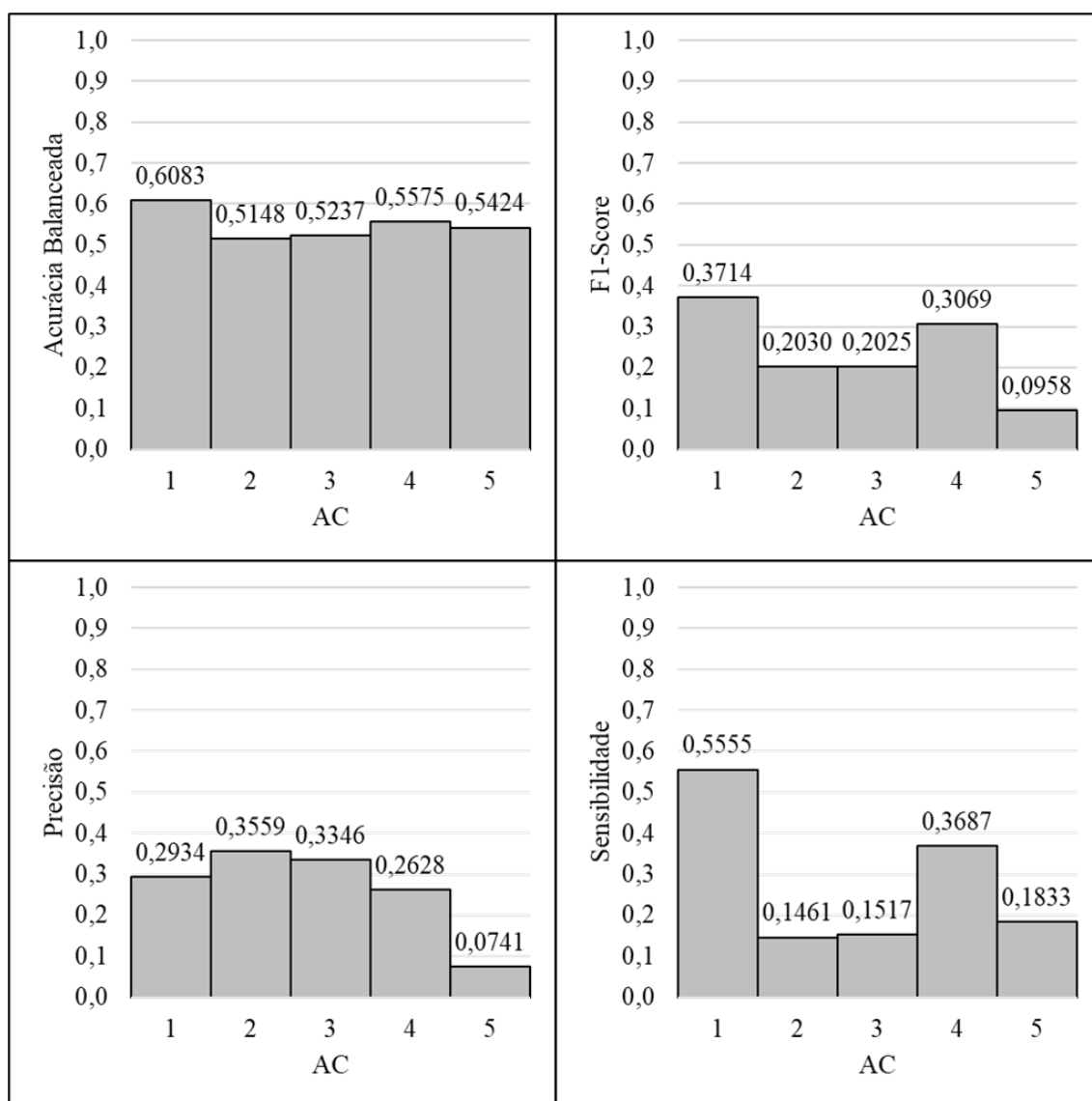


Figura 10 – Médias das estatísticas de desempenho de classe acurácia balanceada, F1-Score, precisão e sensibilidade das parcelas de cada nível de acamamento (AC) dos modelos treinados a partir das alturas do dossel estimadas por análise de imagens como o percentil 50% (MRF_P50). As médias foram calculadas a partir dos valores das estatísticas nos dois modelos de Random Forest (MRF), um treinado com o conjunto de dados do ensaio 1 (E1) e o outro com o conjunto de dados do ensaio (E2). (a) Acurácia balanceada. (b) F1-Score. (c) Precisão. (d) Sensibilidade.

Os níveis 2 e 3 apresentaram as precisões mais altas, mas também apresentaram as sensibilidades mais baixas (Figura 10). Essas estatísticas indicam que, embora parcelas em outros AC não tenham sido classificadas nos níveis 2 e 3, grande parte das parcelas com AC real 2 e 3 foram classificadas em outros níveis. A subestimação do número de parcelas desses níveis, observada na matriz de confusão (Tabela 11), concorda com as conclusões resultantes da análise dessas estatísticas de desempenho.

O nível 5 de AC mostrou o pior desempenho quanto a F1-Score e precisão (Figura 10). A sensibilidade foi consideravelmente inferior às dos níveis 1 e 4 devido ao grande número de falsos positivos gerado pela classificação incorreta de parcelas dos níveis 2 e 3, principalmente no MRF treinado com o conjunto de dados de E1. Ao mesmo tempo, as parcelas pertencentes ao nível 5 foram incorretamente classificadas, em grande parte, nos níveis 1 e 4 de AC.

Embora as parcelas com plantas predominantemente eretas ou levemente acamadas tenham sofrido maior subestimação da ALTDI devido à dificuldade de detecção da extremidade das hastes principais, a acurácia classificatória dos MRFs foi superior em parcelas com AC que indicam menor acamamento. Enquanto as parcelas severamente acamadas podem ter sofrido o tombamento de plantas em diversos períodos ao longo do seu desenvolvimento, parcelas menos acamadas mantiveram um padrão de comportamento da altura do dossel ao longo da série temporal, entre o florescimento e o período anterior ao início da maturação. A modelagem desse padrão pelas curvas ajustadas e a detecção do padrão na porção inicial e mediana da série temporal podem ter favorecido à acurácia preditiva dos MRFs na classificação dessas classes. Além disso, o erro out-of-the-bag, usado para encontrar os valores ótimos dos hiperparâmetros, é suscetível ao desbalanceamento de classes (Janitza & Hornung, 2018). A maior frequência de parcelas nos níveis 1 e 2 de AC, em relação às parcelas nos níveis 4 e 5, pode ter favorecido a seleção de hiperparâmetros eficientes na detecção das parcelas menos acamadas.

A comparação das estatísticas de classe dos modelos que apresentaram o melhor e o pior desempenho, os modelos treinados a partir das alturas do dossel estimadas por análise de imagens como o percentil 50% (MRF_P50) e como o intervalo de altura predominante (MRF_IAP) (Figura 11), respectivamente, revela que o MRF_P50 se destacou principalmente na classificação do nível 1 de AC, com desempenho superior em todas as métricas e especialmente na sensibilidade. As métricas indicam que o MRF_P50 teve um desempenho superior em identificar corretamente e distinguir as parcelas do nível 1, e assim evitou falsos positivos e falsos negativos de forma mais eficaz. O primeiro modelo também se destaca quanto à classificação correta das parcelas do nível 4, cujo desempenho foi consideravelmente superior em todas as estatísticas de classe estimadas.

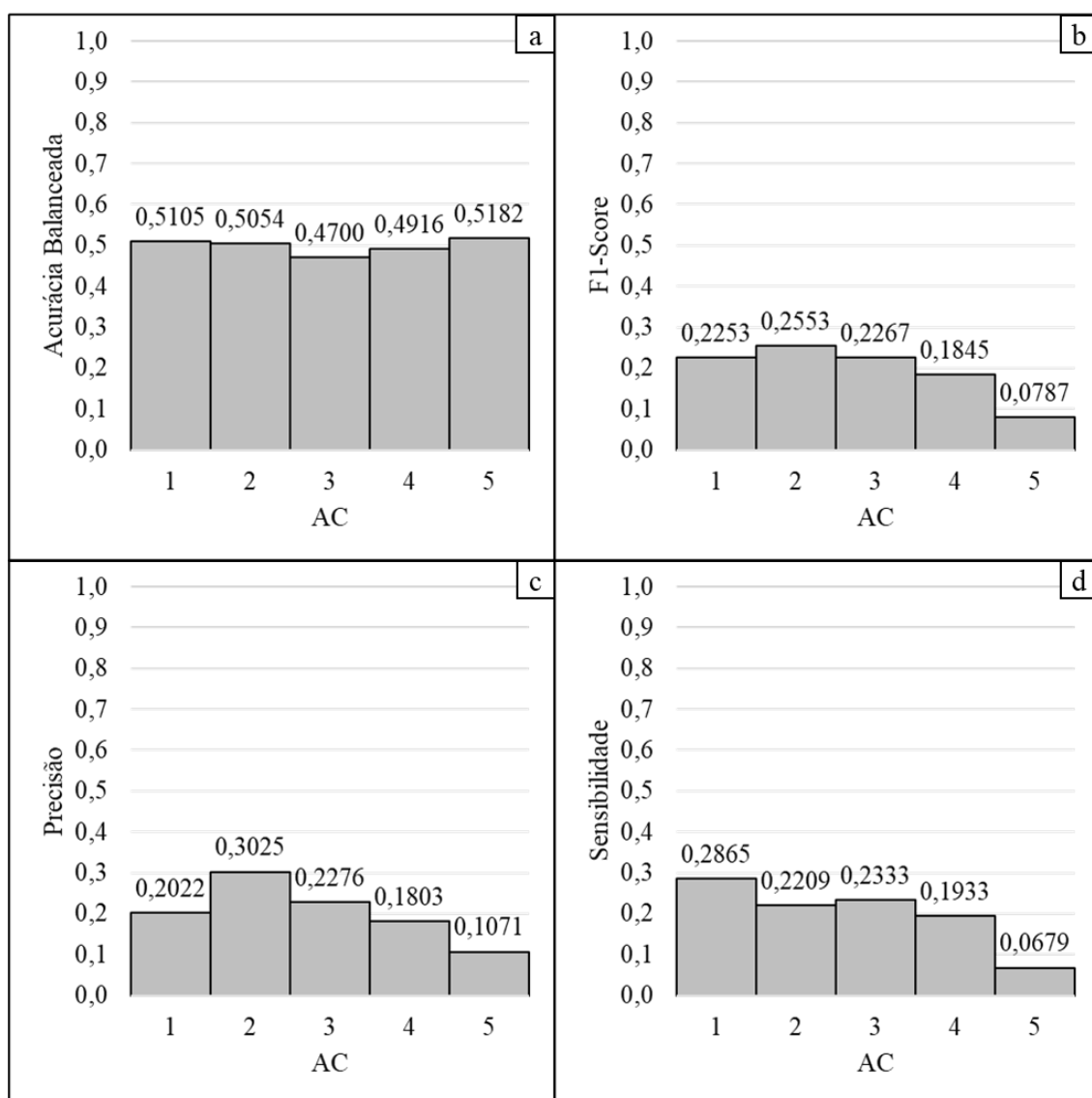


Figura 11 – Médias das estatísticas de desempenho de classe acurácia balanceada, F1-Score, precisão e sensibilidade das parcelas de cada nível de acamamento (AC) dos modelos treinados a partir das alturas do dossel estimadas por análise de imagens como o intervalo de altura predominante (MRF_IAP). As médias foram calculadas a partir dos valores das estatísticas nos dois modelos de Random Forest (MRF), um treinado com o conjunto de dados do ensaio 1 (E1) e o outro com o conjunto de dados do ensaio (E2). (a) Acurácia balanceada. (b) F1-Score. (c) Precisão. (d) Sensibilidade.

Os dois modelos mostraram desempenhos baixos e similares na classificação dos níveis 2, 3 e 5. A precisão de MRF_IAP foi ligeiramente superior à de MRF_P50, e o oposto foi observado na sensibilidade, resultando em F1-Score similares.

O menor desempenho de MRF_IAP em classificar corretamente as parcelas do nível 1 de AC pode estar relacionado à menor representação da extremidade superior das plantas no período anterior à maturação. O método de detecção do intervalo de altura com base na frequência dos valores dos pixels, associado às dificuldades de detecção de

plantas após a redução da área do dossel, levou a uma redução drástica e precoce da altura do dossel normalizadas pelo valor máximo (N_ALTDI) nas parcelas desse nível (Figura 8). Consequentemente, houve maior sobreposição dos parâmetros das curvas ajustadas a parcelas com plantas predominantemente eretas e a com plantas severamente acamadas, o que reduziu a capacidade dos MRFs em as distinguir.

Murakami et al. (2012) estudaram a mensuração quantitativa do acamamento de trigo-sarraceno através da altura do dossel estimada por técnicas de fotogrametria. O acamamento das parcelas ocorreu entre os estádios de florescimento e enchimento de grãos devido à incidência de fortes ventos. A altura do dossel via SfM, avaliada em diferentes estádios de desenvolvimento da cultura, resultou em estimações precisas durante a floração e enchimento de grãos, mas subestimou a altura no estágio de colheita devido à senescência de flores e folhas. As variações observadas ao longo do desenvolvimento da cultura indicam que o modelo pode ser aperfeiçoado. Apesar disso, os autores relataram que o modelo foi capaz de distinguir as cultivares quanto à resistência ao acamamento.

Sarkar et al. (2023) estudaram a aplicação de técnicas de aprendizado de máquinas e ortomosaicos de imagens aéreas RGB para classificar 1266 parcelas de soja em níveis de acamamento. A partir de um voo realizado no estágio de desenvolvimento R6, foram estimadas 12 características de textura de cada parcela. O método SMOTE-ENN foi empregado para tratar o desbalanceamento de dados no pré-processamento. A classificação do acamamento das parcelas foi testada através de quatro algoritmos de aprendizado de máquinas: reforço de gradiente extremo (XGBoost), MRF, k-vizinhos mais próximos e rede neural artificial. Os conjuntos de treinamento e teste foram formados por 80% e 20% dos dados balanceados, respectivamente. A rede neural artificial se mostrou superior aos demais modelos, com acurácia de 96%. De acordo com a matriz de confusão e estatísticas de classe apresentadas, as parcelas não acamadas apresentaram erros de classificação superiores às demais classes em todos os modelos.

A classificação de níveis de acamamento em parcelas de soja baseada em características de textura dos ortomosaicos se mostrou muito superior em relação à baseada em série temporal de ALTDI e ajuste de curva logística. Apesar disso, o primeiro modelo também apresentou dificuldades em classificar corretamente parcelas nas quais não houve acamamento. A acurácia dos modelos favoreceu a observação de que não houve diferenças de desempenho entre os modelos testados, o que indicou que, embora não tenha sido o melhor, o MRF foi adequado para a classificação da severidade do

acamamento em parcelas de soja. Embora bem sucedida, a metodologia foi aplicada a um único estágio de desenvolvimento da soja. Sarkar et al. (2023) não estudaram o desempenho do modelo após a senescência das folhas e, quando aplicada ao estágio R6, a metodologia não capturou informações sobre o acamamento ocorrido no período da maturação.

A predição da severidade do acamamento ocorrido ao longo de todo o desenvolvimento da cultura, independentemente de eventos específicos, exige a análise da curva de desenvolvimento da cultura como um todo. Dessa forma, a subestimação da ALTDI no estágio da maturação compromete o padrão esperado nas séries temporais e reduz a distinção entre os níveis de AC, o que leva a erros classificatórios mais graves quando comparados à detecção do acamamento em períodos específicos que antecedem a senescência das folhas.

3.3.3.3. Detecção do acamamento via classificação dos níveis de acamamento

A utilização dos modelos para a detecção do acamamento das parcelas consistiu em formar o grupo das parcelas com acamamento leve, com as parcelas dos níveis de acamamento (AC) 1 e 2 e relação esperada entre o comprimento de plantas na maturação (COMPM) e altura do dossel avaliada no campo (ALTDC) superior a 0,809, e o grupo com acamamento severo, composto pelas parcelas dos demais AC e relação esperada entre COMPM e ALTDC inferior a 0,809.

A matriz de confusão (Tabela 13) foi construída por meio do agrupamento das parcelas de acordo com os AC reais e preditos na classificação executada para a predição dos cinco níveis de AC. Os modelos treinados a partir das alturas do dossel estimadas por análise de imagens como o percentil 50% (MRF_P50) apresentaram o maior número médio de classificações corretas dentre os modelos de Random Forest (MRF) avaliados, mas foram inconsistentes quando treinados com conjuntos de dados diferentes. Enquanto o treinamento com o conjunto de dados do ensaio 1 (E1) cometeu um maior número de erros ao classificar incorretamente parcelas com acamamento leve como severo, o oposto ocorreu quando o treinamento foi realizado com o conjunto de dados do ensaio 2 (E2).

Tabela 13 – Matrizes de confusão formadas pelo número de parcelas de cada nível de acamamento real (AC Real) e seus respectivos níveis de acamamento preditos (AC Predito) pelos modelos de Random Forest (MRF) de cada método de estimação da altura do dossel por análise de imagens (ME_ALTDI) utilizados para a detecção do acamamento. Os níveis 1 e 2 de AC foram agrupados para formar a classe “acamamento leve” e os níveis 3, 4 e 5 foram agrupados para formar a classe “acamamento severo”. As colunas à esquerda representam o resultado do treinamento e teste realizados com os conjuntos de dados dos ensaios 1 (E1) e 2 (E2), respectivamente, enquanto as à direita representam o resultado do treinamento e teste com E2 e E1, respectivamente.

			Treinamento com E1 e teste com E2			Treinamento com E2 e teste com E1		
			AC Real			AC Real		
			Leve	Severo	Total	Leve	Severo	Total
MRF_P50	AC	Leve	96	46	142	88	99	187
	Predito	Severo	72	92	164	58	96	154
MRF_P75	AC	Leve	93	58	151	92	109	201
	Predito	Severo	75	80	155	54	86	140
MRF_P90	AC	Leve	91	71	162	86	104	190
	Predito	Severo	77	67	144	60	91	151
MRF_P95	AC	Leve	86	57	143	94	128	222
	Predito	Severo	82	81	163	52	67	119
MRF_P99	AC	Leve	78	53	131	82	64	146
	Predito	Severo	90	85	175	86	109	195
MRF_IAP	AC	Leve	72	55	127	87	101	188
	Predito	Severo	96	83	179	59	94	153
Total			168	138	306	146	195	341

MRF_P50: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 50%; MRF_P75: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 75%; MRF_P90: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 90%; MRF_P95: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 95%; MRF_P99: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 99%; MRF_IAP: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o intervalo de altura predominante.

A mesma tendência foi observada nos demais MRF. Embora o número de verdadeiros positivos tenha sido similar, quando treinados com o conjunto de dados de E1, os MRFs cometeram mais erros do tipo falso negativo, enquanto com E2, houve mais falsos positivos. Apesar de terem desempenhos consistentes quando treinados com diferentes conjuntos de dados, os MRF comumente cometem erros classificatórios quando as variáveis preditoras de diferentes classes têm valores similares ou sobrepostos (Fernandez-Delgado et al., 2014; Rawat & Mishra, 2022).

As estatísticas de desempenho geral dos MRFs para a detecção do acamamento indicam que o MRF_P50 foi, em média, superior aos demais métodos em todas as métricas. Apesar disso, o desempenho do modelo foi drasticamente diferente quando os

conjuntos de treinamento e teste foram invertidos. O desempenho do MRF_P50 treinado com o conjunto de dados de E1 foi muito superior quando comparado ao treinado com E2. Enquanto o MRF_P50 foi superior a todos os outros MRFs quando o treinamento foi realizado com o conjunto de dados de E1, sua acurácia balanceada e coeficiente Kappa foram inferiores, e seu F1-Score foi equivalente ao de MRF_P99 quando o treinamento foi realizado com E2 (Tabela 14).

Tabela 14 – Estatísticas de desempenho geral acurácia balanceada, F1-Score e coeficiente Kappa dos modelos de Random Forest (MRF) utilizados na detecção do acamamento, com conjuntos de dados de treinamento E1 e E2. Os níveis 1 e 2 de AC foram agrupados para formar a classe “acamamento leve”, enquanto os níveis 3, 4 e 5 foram agrupados para formar a classe “acamamento severo”.

FA	Treinamento com E1 e teste com E2			Treinamento com E2 e teste com E1		
	Acurácia Balanceada	F1- Score	Coeficiente Kappa	Acurácia Balanceada	F1- Score	Coeficiente Kappa
MRF_P50	0,6190	0,6194	0,2342	0,5475	0,5285	0,0918
MRF_P75	0,5666	0,5831	0,1318	0,5356	0,5303	0,0679
MRF_P90	0,5136	0,5515	0,0271	0,5279	0,5119	0,0537
MRF_P95	0,5494	0,5531	0,0973	0,4937	0,5109	0,0000
MRF_P99	0,5401	0,5217	0,0784	0,5603	0,5223	0,1184
MRF_IAP	0,5150	0,4881	0,0292	0,5390	0,5210	0,0752
Média	0,5506	0,5528	0,0997	0,5340	0,5208	0,0678

MRF_P50: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 50%; MRF_P75: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 75%; MRF_P90: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 90%; MRF_P95: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 95%; MRF_P99: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 99%; MRF_IAP: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o intervalo de altura predominante.

A diferença de performance do MRF_P50 quando treinado com diferentes conjuntos de dados pode ser explicada pelo número de parcelas e pela proporção entre classes de E1 e E2. O E1 dispunha de 168 parcelas levemente acamadas e 146 severamente acamadas, enquanto E2 dispunha de 138 parcelas levemente acamadas e 195 parcelas severamente acamadas, o que resulta em razões entre parcelas levemente acamadas e severamente acamadas iguais a 1,15 e 0,71, respectivamente.

Embora seu impacto não seja tão drástico quanto em outros modelos de aprendizado de máquinas, o aumento do tamanho populacional do conjunto de treinamento geralmente aumenta a acurácia preditiva de MRF. Em conjuntos de dados de treinamento pequenos, com cerca de 500 amostras, o equilíbrio entre classes no conjunto

de treinamento é especialmente importante para a acurácia dos modelos (Liu et al., 2023; Ramezan et al., 2021).

Schulthess et al. (2023) estudaram o efeito do tamanho amostral e da proporção das amostras na classificação de nove culturas agrícolas através de MRF e análise de imagens de satélite. Os autores reportam que o aumento do número de amostras no conjunto de treinamento aumenta a acurácia dos modelos e que a representação equilibrada das classes resulta em uma maior acurácia preditiva.

O maior número de parcelas com acamamento leve no conjunto de treinamento pode ser responsável pelo menor número de falsos positivos e pela tendência de superestimação do número de dessa classe quando o treinamento foi realizado com o conjunto de dados do ensaio 1 (E1). No modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 50% (MRF_P50), o efeito do maior equilíbrio entre o número de parcelas de cada classe e maior número de parcelas totais pode ser responsável pela melhor capacidade preditiva do MRF treinado com o conjunto de dados de E1, em comparação treinado com o ensaio 2 (E2).

Ao serem comparadas com a aplicação dos MRFs para classificação dos níveis de AC, a acurácia balanceada e o coeficiente Kappa (Tabela 14) mostram que, com exceção de MRF_P50 e MRF_P75 treinados com o conjunto de dados de E1, nenhum dos MRFs apresentou ganho de desempenho significativo ao ser utilizado na detecção do acamamento. Esse resultado indica que os erros classificatórios não se limitaram apenas aos AC semelhantes quanto à severidade do acamamento, mas se estenderam a todos os níveis de AC.

Acorsi et al. (2019) avaliaram o potencial do uso de VANT para a detecção do acamamento através de um único voo, realizado em um campo de produção de milho no estádio de enchimento de grãos, após fortes chuvas e ventos danificarem a lavoura. A metodologia utilizada foi baseada em um MDE obtido do processamento de imagens aéreas RGB e SfM. Regiões de interesse foram selecionadas aleatoriamente no campo, e as áreas acamadas dessas regiões foram identificadas visualmente através do ortomosaico. A altura do dossel de cada região de interesse foi definida como o P99 máximo observado. Os limites de 80%, 70%, 60% e 50% de redução da altura do dossel foram testados quanto à acurácia em determinar áreas acamadas. O limite de 60% resultou em um coeficiente de correlação de 0,93 e RMSE% de 8,72% entre a área acamada determinada visualmente e através do limite de redução da altura.

Embora simples e eficiente, a metodologia avaliada pelos autores apresenta limitações ao ser aplicada a ensaios de melhoramento devido às variações de altura do dossel inerentes à diversidade genética do material cultivado em cada parcela. A aplicação da metodologia a séries temporais de altura do dossel de parcelas de soja incorreria nos mesmos erros provocados por variações ligadas à maturação das plantas e subestimação da ALTDI de parcelas levemente acamadas.

3.3.3.4. Considerações sobre o desempenho dos MRFs treinados

As estatísticas de desempenho geral dos modelos de Random Forest (MRF) treinados com os conjuntos de dados dos ensaios 1 (E1) e 2 (E2) (Tabela 12 e Tabela 14) sugerem que o modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 50% (MRF_P50) foi superior aos demais MRFs na predição do nível do acamamento (AC) e na detecção do acamamento, segundo a metodologia testada. Os valores das acurácias balanceada, F1-Score e coeficientes Kappa sugerem que, embora não tenham sido suficientemente caracterizados e identificados pelos MRFs nos conjuntos de dados testados, existem padrões nos parâmetros das curvas ajustadas à série temporal de alturas do dossel normalizadas pelo valor máximo (N_ALTDI) que permitem identificar a ocorrência e prever o nível de severidade do acamamento de parcelas experimentais de soja.

A inspeção das etapas do processamento de imagens (Figura 12) fornece evidências de que a subestimação das alturas das extremidades das hastes principais nos MDEs dos dosséis, no período que antecedeu a maturação das parcelas predominantemente eretas e levemente acamadas foi prejudicial à estimação da altura do dossel por análise de imagens (ALTDI). A senescência e queda de folhas levaram à suavização das reconstruções tridimensionais de parcelas com acamamento leve (Figura 12c), enquanto a retenção de folhas e hastes verdes por parte das parcelas com acamamento severo levaram à superestimação das ALTDI de parcelas com acamamento severo.

A subestimação das alturas do dossel normalizadas pelo valor máximo (N_ALTDI) de parcelas com acamamento leve e relação esperada entre o comprimento de plantas na maturação (COMPM) e altura do dossel avaliada no campo (ALTDC) alta e superestimação das N_ALTDI de parcelas com acamamento severo e relação esperada entre COMPM e ALTDC baixas, no final do ciclo, impediu a seleção de um método de estimação da altura do dossel por análise de imagens (ME_ALTDI) capaz de representar

adequadamente as alturas dos dosséis de parcelas com diferentes níveis de acamamento (AC) e em diferentes estádios de desenvolvimento. A baixa acurácia da estimação das N_ALTDI no período anterior à maturação levou à construção de séries temporais e ajuste de curvas logísticas semelhantes em parcelas com diferentes AC, o que limitou a distinção entre os parâmetros das curvas ajustadas e, conseqüentemente, a capacidade dos MRF de capturar o comportamento divergente das alturas de plantas esperado e observado no campo.

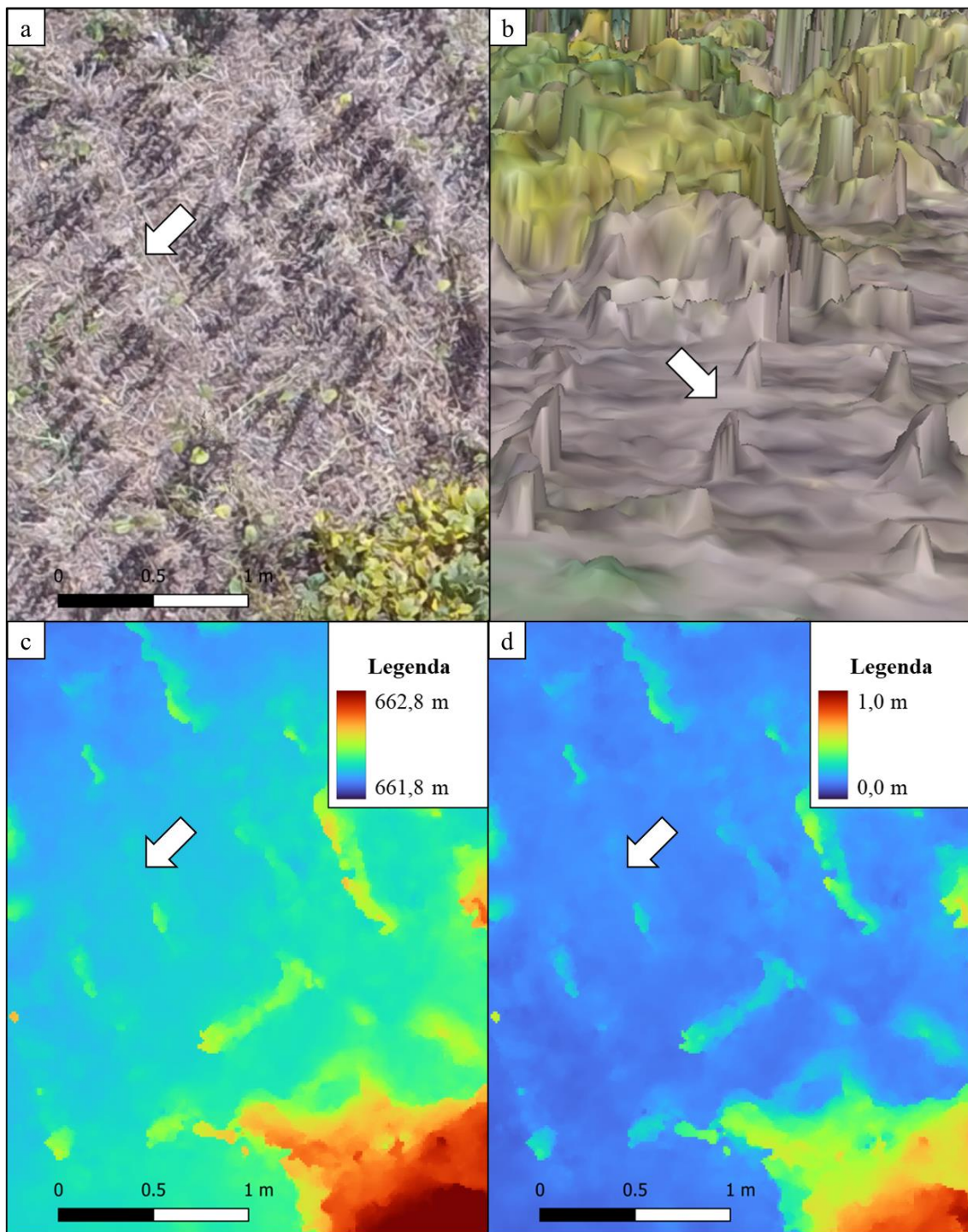


Figura 12 – Etapas do processamento de imagens do voo 14. As setas brancas indicam uma parcela com plantas eretas em data próxima à sua maturação. (a) Ortomosaico gerado a partir das imagens RGB. (b) Malha gerada a partir da reconstrução tridimensional da nuvem densa de pontos. (c) Modelo digital de elevação (MDE) gerado por *Surface from motion* (SfM). (d) Modelo digital de altura dos dosséis obtido da subtração do MDE do solo do MDE original.

A Figura 12a destaca uma área de interesse com parcelas eretas após a senescência e queda das folhas, devido à maturação. A presença das plantas, seu ângulo de inclinação e suas alturas são evidenciadas pela presença de sombras no solo. A Figura

12b mostra a mesma área representada na reconstrução tridimensional do cenário, a partir da nuvem de pontos. A comparação da parcela indicada com plantas cujas folhas ainda estão verdes, ao fundo da imagem, mostra a subestimação da altura da parcela. Além disso, é possível notar que algumas plantas não foram detectadas na reconstrução do cenário. As Figura 12c e Figura 12d mostram a área de interesse no modelo digital de elevação original e o MDE dos dosséis, em que é possível notar a subestimação das alturas e a não detecção de algumas plantas, como as diretamente indicadas pela seta.

A reconstrução de cenários tridimensionais a partir de SfM e imagens RGB é baseada apenas em técnicas de análise de imagens, o que pode gerar erros e baixa acurácia, especialmente em cenários complexos. Essa metodologia apresenta limitações em cenários com vegetação de altura mediana, visto que as plantas podem bloquear a linha de visão e limitar o número de imagens que representam os detalhes de plantas adjacentes (Mandici & Nedevschi, 2015).

Sensores LiDAR, que estimam a altimetria do cenário através da emissão de lasers e mensuração do intervalo de tempo até o retorno do sinal ao sensor, são capazes de gerar modelos tridimensionais com alta precisão quanto à elevação e ao detalhamento (Journal, 2022; Octariady et al., 2017). Por outro lado, MDE gerados através de SfM e imagens RGB geralmente apresentam limitações em capturar detalhes finos de forma precisa (Mandici & Nedevschi, 2015).

Além disso, sensores LiDAR são capazes de penetrar na folhagem e detectar objetos ocultos pelo dossel das plantas, o que torna possível detectar a estrutura da haste principal, e assim, eliminar os erros causados pela retenção de folhas verdes em plantas acamadas e pelas folhas da parte superior do dossel, ao estimar altura das plantas no período anterior à maturação (Tulldahl et al., 2019).

Salach et al. (2018) estudaram a acurácia vertical de MDE gerados através de dados de LiDAR e de Dense Image Matching (DIM), uma técnica de análise de imagens similar ao SfM. Os autores reportaram que o LiDAR ofereceu melhor resolução espacial e densidade da nuvem densa de pontos. Para áreas com vegetação de altura média, de cerca de 60 cm, o LiDAR e o DIM resultaram em RMSE de 0,11 m e 0,36 m, respectivamente, o que indica a superioridade do LiDAR na estimação da altura de plantas nesse tipo de cenário. Os autores concluíram que o LiDAR apresentou maior acurácia vertical comparado ao MDE gerado por DIM.

O uso de sensores LiDAR acoplados a VANT é uma alternativa promissora para a predição do AC através de séries temporais de alturas de plantas. A maior acurácia

vertical, detalhamento do MDE e capacidade de detectar objetos ocultos pela folhagem podem superar as limitações impostas pela redução da área do dossel devido à queda de folhas. A detecção direta das hastes principais e dos racemos terminais pode reduzir os erros de estimação das ALTDI e fazer com que as séries temporais e curvas logísticas ajustadas representem o comportamento da altura do dossel de forma mais robusta e consistente com as alturas observadas no campo.

4. CONCLUSÕES

Neste trabalho, foi avaliado o potencial da utilização de imagens aéreas RGB e de modelos digitais de elevação reconstruídos através de *Surface from motion* na estimação da altura do dossel de parcelas de soja com diferentes níveis de acamamento. A série de voos executada possibilitou a análise de imagens e a construção de modelos digitais de elevação com alta resolução espacial e temporal.

A estimação da altura do dossel a partir de imagens aéreas RGB e *Surface from motion* no estágio do florescimento como o percentil 50% resultou em acurácia satisfatória, com raiz do erro quadrático médio e raiz do erro quadrático médio percentual iguais a 12,9 cm e 13,5%, respectivamente. No estágio da maturação, o percentil 99% resultou em acurácia moderada, com raiz do erro quadrático médio e raiz do erro quadrático médio percentual iguais a 28,8 cm e 18,7%, respectivamente, mas todos os métodos de estimação por análise de imagens mostraram precisão limitada na estimação da altura do dossel parcelas predominantemente eretas e drasticamente acamadas.

O modelo de Random Forest treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 50% mostraram a existência de padrões distintos nos parâmetros das curvas ajustadas à série temporal de alturas do dossel normalizadas pelo valor máximo dos diferentes níveis de acamamento, mas as limitações na estimação acurada da altura dos dosséis no período que antecede a maturação reduziram o potencial do modelo em classificar as parcelas quanto à severidade e ocorrência do acamamento.

Inferências sobre o acamamento através de informações obtidas por sensoriamento remoto e diretamente associadas à evolução da altura das plantas apresentam grande potencial e versatilidade de uso no melhoramento genético. A utilização da metodologia testada pode ter maior potencial de uso se aplicada a modelos digitais de elevação estimados através de sensores capazes de representar as alturas das plantas de forma mais acurada e detalhada, e que não sofram interferências pela queda ou retenção de folhas no final do ciclo da cultura.

REFERÊNCIAS

- Abebe, A. M., Kim, Y., Kim, J., Kim, S. L., & Baek, J. (2023). Image-Based High-Throughput Phenotyping in Horticultural Crops. *Plants*, 12(10), 2061. <https://doi.org/10.3390/plants12102061>
- Acorsi, M. G., Martello, M., & Angnes, G. (2019). Identification of Maize Lodging: A Case Study Using a Remotely Piloted Aircraft System. *Engenharia Agrícola*, 39, 66–73. <https://doi.org/10.1590/1809-4430-Eng.Agric.v39nep66-73/2019>
- Alliprandini, L. F., Abatti, C., Bertagnolli, P. F., Cavassim, J. E., Gabe, H. L., Kurek, A., Matsumoto, M. N., De Oliveira, M. A. R., Pitol, C., Prado, L. C., & Steckling, C. (2009). Understanding Soybean Maturity Groups in Brazil: Environment, Cultivar Classification, and Stability. *Crop Science*, 49(3), 801–808. <https://doi.org/10.2135/cropsci2008.07.0390>
- Anand, K., & Prashant, K. (2023). A Review on High Throughput Phenotyping for Vegetable Crops. *Journal of Botany Research*, 6(1). <https://doi.org/10.36959/771/575>
- Bai, D., Li, D., Zhao, C., Wang, Z., Shao, M., Guo, B., Liu, Y., Wang, Q., Li, J., Guo, S., Wang, R., Li, Y., Qiu, L., & Jin, X. (2022). Estimation of soybean yield parameters under lodging conditions using RGB information from unmanned aerial vehicles. *Frontiers in Plant Science*, 13, 1012293. <https://doi.org/10.3389/fpls.2022.1012293>
- Bakhtawer, B., Naz, B., Narejo, S., Din, N. U., & Ahmed, W. (2022). Soil Classification & Prediction of Crop Status with Supervised Learning Algorithm: Random Forest. *International Journal of Innovations in Science and Technology*, 4(4), 1011–1022. <https://doi.org/10.33411/IJIST/2022040403>
- Bakirarar, B., & Elhan, A. H. (2023). Class Weighting Technique to Deal with Imbalanced Class Problem in Machine Learning: Methodological Research. *Turkiye Klinikleri Journal of Biostatistics*, 15(1), 19–29. <https://doi.org/10.5336/biostatic.2022-93961>
- Bernardelli De Moraes, M., & Palermo Coelho, G. (2023). Effects of the Random Forests Hyper-Parameters in Surrogate Models for Multi-Objective Combinatorial Optimization: A Case Study using MOEA/D-RFTS. *IEEE Latin America Transactions*, 21(5), 621–627. <https://doi.org/10.1109/TLA.2023.10130833>
- Bhandari, M., Chang, A., Jung, J., Ibrahim, A. M. H., Rudd, J. C., Baker, S., Landivar, J., Liu, S., & Landivar, J. (2023). Unmanned aerial system-based high-throughput

- phenotyping for plant breeding. *The Plant Phenome Journal*, 6(1), e20058. <https://doi.org/10.1002/ppj2.20058>
- Bhat, J. A., Karikari, B., Adeboye, K. A., Ganie, S. A., Barmukh, R., Hu, D., Varshney, R. K., & Yu, D. (2022). Identification of superior haplotypes in a diverse natural population for breeding desirable plant height in soybean. *Theoretical and Applied Genetics*, 135(7), 2407–2422. <https://doi.org/10.1007/s00122-022-04120-0>
- Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A., Deng, D., & Lindauer, M. (2023). Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. *WIREs Data Mining and Knowledge Discovery*, 13(2), e1484. <https://doi.org/10.1002/widm.1484>
- Borra-Serrano, I., De Swaef, T., Quataert, P., Aper, J., Saleem, A., Saeys, W., Somers, B., Roldán-Ruiz, I., & Lootens, P. (2020). Closing the Phenotyping Gap: High Resolution UAV Time Series for Soybean Growth Analysis Provides Objective Data from Field Trials. *Remote Sensing*, 12(10), 1644. <https://doi.org/10.3390/rs12101644>
- Buelvas, R. M., Adamchuk, V. I., Lan, J., Hoyos-Villegas, V., Whitmore, A., & Stromvik, M. V. (2023). Development of a Quick-Install Rapid Phenotyping System. *Sensors*, 23(9), 4253. <https://doi.org/10.3390/s23094253>
- Casella, E., Lewin, P., Ghilardi, M., Rovere, A., & Bejarano, S. (2022). Assessing the relative accuracy of coral heights reconstructed from drones and structure from motion photogrammetry on coral reefs. *Coral Reefs*, 41(4), 869–875. <https://doi.org/10.1007/s00338-022-02244-9>
- Chang, A., Jung, J., Maeda, M. M., & Landivar, J. (2017). Crop height monitoring with digital imagery from Unmanned Aerial System (UAS). *Computers and Electronics in Agriculture*, 141, 232–237. <https://doi.org/10.1016/j.compag.2017.07.008>
- Chen, H., Yang, Z., Chen, L., Zhang, C., Yuan, S., Zhang, X., Qiu, D., Wan, Q., Zhan, Y., Chen, S., Shan, Z., & Zhou, X. (2017). Combining QTL and candidate gene analysis with phenotypic model to unravel the relationship between lodging and related traits in soybean. *Molecular Breeding*, 37(4), 43. <https://doi.org/10.1007/s11032-017-0645-5>

- Chu, T., Starek, M. J., Brewer, M. J., Murray, S. C., & Pruter, L. S. (2018). Characterizing canopy height with UAS structure-from-motion photogrammetry—Results analysis of a maize field trial with respect to multiple factors. *Remote Sensing Letters*, 9(8), 753–762. <https://doi.org/10.1080/2150704X.2018.1475771>
- Cober, E. R., & Morrison, M. J. (2022). Precipitation Irregularity and Solar Radiation Play a Role in Determining Short-Season Soybean Yield. *Canadian Journal of Plant Science*, CJPS-2022-0104. <https://doi.org/10.1139/CJPS-2022-0104>
- Devi, V. B., Prabavathi, R., Subha, P., & Meenaloshini, M. (2022). An Efficient and Robust Random Forest Algorithm for Crop Disease Detection. *2022 International Conference on Communication, Computing and Internet of Things (IC3IoT)*, 1–4. <https://doi.org/10.1109/IC3IoT53935.2022.9767937>
- Dhillon, M. S., Dahms, T., Kuebert-Flock, C., Rummler, T., Arnault, J., Steffan-Dewenter, I., & Ullmann, T. (2023). Integrating random forest and crop modeling improves the crop yield prediction of winter wheat and oil seed rape. *Frontiers in Remote Sensing*, 3, 1010978. <https://doi.org/10.3389/frsen.2022.1010978>
- Elzhov, T. V., Mullen, K. M., Spiess, A.-N., Bolker, B., Mullen, M. K. M., & Suggests, M. (2016). Package ‘minpack. Lm’. Title *R Interface Levenberg-Marquardt Nonlinear Least-Sq. Algorithm Found MINPACK Plus Support Bounds*. <ftp://r-project.org/pub/R/web/packages/minpack.lm/minpack.lm.pdf>
- Fan, L. (2020). A Comparison Between Structure-From-Motion and Terrestrial Laser Scanning for Deriving Surface Roughness: A Case Study on a Sandy Terrain Surface. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-3/W10, 1225–1229. <https://doi.org/10.5194/isprs-archives-XLII-3-W10-1225-2020>
- Fernandez-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? *Journal of Machine Learning Research*, 15(1), 3133–3181.
- Florez, A. Y. C., Vines, B. V. S., Arroyo, D. C. R., Saire, J. E. C., & Franco, P. B. (2020). Experimental Evaluation of Accuracy of Most Common Machine Learning Models using Pulsar Data Set. *2020 International Conference of Digital Transformation and Innovation Technology (Incodtrin)*, 80–83. <https://doi.org/10.1109/Incodtrin51881.2020.00027>
- Govaichelvan, K., Pathmanathan, D., Zainal-Abidin, R., & Abu, A. (2024). Machine learning for major food crops breeding: Applications, challenges, and ways

- forward. *Agronomy Journal*, 116(3), 1112–1125.
<https://doi.org/10.1002/agj2.21393>
- Grüner, E., Astor, T., & Wachendorf, M. (2019). Biomass Prediction of Heterogeneous Temperate Grasslands Using an SfM Approach Based on UAV Imaging. *Agronomy*, 9(2), 54. <https://doi.org/10.3390/agronomy9020054>
- Hairani, H., Anggrawan, A., & Priyanto, D. (2023). Improvement Performance of the Random Forest Method on Unbalanced Diabetes Data Classification Using Smote-Tomek Link. *JOIV: International Journal on Informatics Visualization*, 7(1), 258. <https://doi.org/10.30630/joiv.7.1.1069>
- Han, X., Thomasson, J. A., Bagnall, C., Rooney, W. L., Pugh, N. A., Horne, D. W., Malambo, L., Jung, J., Chang, A., & Cope, D. A. (2018). Calibrated plant height estimates with structure from motion from fixed-wing UAV images. In J. A. Thomasson, M. McKee, & R. J. Moorhead (Orgs.), *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping III* (p. 11). SPIE. <https://doi.org/10.1117/12.2305746>
- Hand, D. J., Christen, P., & Kirielle, N. (2021). F*: An interpretable transformation of the F-measure. *Machine Learning*, 110(3), 451–456. <https://doi.org/10.1007/s10994-021-05964-1>
- Hassani, K., Gholizadeh, H., Taghvaeian, S., Natalie, V., Carpenter, J., & Jacob, J. (2023). Application of UAS-Based Remote Sensing in Estimating Winter Wheat Phenotypic Traits and Yield During the Growing Season. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*. <https://doi.org/10.1007/s41064-022-00229-5>
- Herr, A. W., Adak, A., Carroll, M. E., Elango, D., Kar, S., Li, C., Jones, S. E., Carter, A. H., Murray, S. C., Paterson, A., Sankaran, S., Singh, A., & Singh, A. K. (2023). Unoccupied aerial systems imagery for phenotyping in cotton, maize, soybean, and wheat breeding. *Crop Science*, 63(4), 1722–1749. <https://doi.org/10.1002/csc2.21028>
- Huang, S., Gao, Y., Li, Y., Xu, L., Tao, H., & Wang, P. (2017). Influence of plant architecture on maize physiology and yield in the Heilonggang River valley. *The Crop Journal*, 5(1), 52–62. <https://doi.org/10.1016/j.cj.2016.06.018>
- Ibiev, G. Z., Savoskina, O. A., Chebanenko, S. I., Beloshapkina, O. O., & Zaverkin, I. A. (2022). *Unmanned aerial vehicles (UAVs)—One of the digitalization and*

- effective development segments of agricultural production in modern conditions.* 070007. <https://doi.org/10.1063/5.0107373>
- Janitza, S., & Hornung, R. (2018). On the overestimation of random forest's out-of-bag error. *PLOS ONE*, 13(8), e0201904. <https://doi.org/10.1371/journal.pone.0201904>
- Jian, Z., Lu, Z., Zhou, X., Lan, B., Xiao, A., Wang, X., & Liang, B. (2022). PUTN: A Plane-fitting based Uneven Terrain Navigation Framework. *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 7160–7166. <https://doi.org/10.1109/IROS47612.2022.9981038>
- Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, 6(1), 27. <https://doi.org/10.1186/s40537-019-0192-5>
- Journal, I. (2022). Design and Implementation of LIDAR System for Distance Measurements. *Interantional Journal of Scientific Research in Engineering and Management*, 06(10). <https://doi.org/10.55041/IJSREM16626>
- Kalimuthu, H., Tan, W. N., Lim, S. L., & Fauzi, M. F. A. (2016). Interpolation of low resolution Digital Elevation Models: A comparison. *2016 8th Computer Science and Electronic Engineering (CEECE)*, 71–76. <https://doi.org/10.1109/CEECE.2016.7835891>
- Kato, S., Samanfar, B., Morrison, M. J., Bekele, W. A., Torkamaneh, D., Rajcan, I., O'Donoghue, L., Belzile, F., & Cober, E. R. (2021). Genome-wide association study to identify soybean stem pushing resistance and lodging resistance loci. *Canadian Journal of Plant Science*, 101(5), 663–670. <https://doi.org/10.1139/cjps-2020-0187>
- Kehl, K., Carvalho, I. R., Sacon, D., Rizzardi, M. A., Langaro, N. C., Loro, M. V., Moura, N. B., & Lautenchleger, F. (2022). Strategic positioning of soybean based on the agronomic ideotype and on fixed and mixed multivariate models. *Pesquisa Agropecuária Brasileira*, 57, e02702. <https://doi.org/10.1590/s1678-3921.pab2022.v57.02702>
- Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*, 28(5). <https://doi.org/10.18637/jss.v028.i05>
- Lee, S., Jun, T. H., Michel, A. P., & Rouf Mian, M. A. (2015). SNP markers linked to QTL conditioning plant height, lodging, and maturity in soybean. *Euphytica*, 203(3), 521–532. <https://doi.org/10.1007/s10681-014-1252-8>

- Li, F., Zhu, H., Luo, Z., Shen, H., & Li, L. (2021). An Adaptive Surface Interpolation Filter Using Cloth Simulation and Relief Amplitude for Airborne Laser Scanning Data. *Remote Sensing*, 13(15), 2938. <https://doi.org/10.3390/rs13152938>
- Liaw, A., & Wiener, M. (2002). *Classification and Regression by randomForest*. 2(3), 18–22.
- Lin, F., Chhapekar, S. S., Vieira, C. C., Da Silva, M. P., Rojas, A., Lee, D., Liu, N., Pardo, E. M., Lee, Y.-C., Dong, Z., Pinheiro, J. B., Ploper, L. D., Rupe, J., Chen, P., Wang, D., & Nguyen, H. T. (2022). Breeding for disease resistance in soybean: A global perspective. *Theoretical and Applied Genetics*, 135(11), 3773–3872. <https://doi.org/10.1007/s00122-022-04101-3>
- Ling, Q. (2023). Machine learning algorithms review. *Applied and Computational Engineering*, 4(1), 91–98. <https://doi.org/10.54254/2755-2721/4/20230355>
- Liu, Y., Zhang, J., P. Ward, M., Tu, W., Yu, L., Shi, J., Hu, Y., Gao, F., Cao, Z., & Zhang, Z. (2023). Impacts of sample ratio and size on the performance of random forest model to predict the potential distribution of snail habitats. *Geospatial Health*, 18(2). <https://doi.org/10.4081/gh.2023.1151>
- Liu, Z., Li, H., Fan, X., Huang, W., Yang, J., Zheng, Y., Wen, Z., Li, Y., Wang, D., Wang, S., & Qiu, L. (2017). Selection of soybean elite cultivars based on phenotypic and genomic characters related to lodging tolerance. *Plant Breeding*, 136(4), 526–538. <https://doi.org/10.1111/pbr.12495>
- Maimaitijiang, M., Sagan, V., Sidike, P., Maimaitiyiming, M., Hartling, S., Peterson, K. T., Maw, M. J. W., Shakoor, N., Mockler, T., & Fritschi, F. B. (2019). Vegetation Index Weighted Canopy Volume Model (CVMVI) for soybean biomass estimation from Unmanned Aerial System-based RGB imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 151, 27–41. <https://doi.org/10.1016/j.isprsjprs.2019.03.003>
- Malambo, L., Popescu, S. C., Murray, S. C., Putman, E., Pugh, N. A., Horne, D. W., Richardson, G., Sheridan, R., Rooney, W. L., Avant, R., Vidrine, M., McCutchen, B., Baltensperger, D., & Bishop, M. (2018). Multitemporal field-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. *International Journal of Applied Earth Observation and Geoinformation*, 64, 31–42. <https://doi.org/10.1016/j.jag.2017.08.014>

- Malhotra, S., & Karanickolas, J. (2020). *A Numerical Transform of Random Forest Regressors corrects Systematically-Biased Predictions* (Versão 1). arXiv. <https://doi.org/10.48550/ARXIV.2003.07445>
- Mandici, S., & Nedevschi, S. (2015). Probabilistic inverse sensor model based Digital Elevation Map creation for an omnidirectional stereovision system. *2015 IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, 231–237. <https://doi.org/10.1109/ICCP.2015.7312635>
- Meyer, M. C., Favoreto, L., Klepker, D., & Marcelino-Guimarães, F. C. (2017). Soybean green stem and foliar retention syndrome caused by *Aphelenchoides besseyi*. *Tropical Plant Pathology*, 42(5), 403–409. <https://doi.org/10.1007/s40858-017-0167-z>
- Millard, K., & Richardson, M. (2015). On the Importance of Training Data Sample Selection in Random Forest Image Classification: A Case Study in Peatland Ecosystem Mapping. *Remote Sensing*, 7(7), 8489–8515. <https://doi.org/10.3390/rs70708489>
- Morrow, J., Willenburg, K. T., & Liscum, E. (2018). Phototropism in land plants: Molecules and mechanism from light perception to response. *Frontiers in Biology*, 13(5), 342–357. <https://doi.org/10.1007/s11515-018-1518-y>
- Murakami, T., Yui, M., & Amaha, K. (2012). Canopy height measurement by photogrammetric analysis of aerial images: Application to buckwheat (*Fagopyrum esculentum* Moench) lodging evaluation. *Computers and Electronics in Agriculture*, 89, 70–75. <https://doi.org/10.1016/j.compag.2012.08.003>
- Najafabadi, M., Hesami, M., & Eskandari, M. (2023). Machine Learning-Assisted Approaches in Modernized Plant Breeding Programs. *Genes*, 14(4), 777. <https://doi.org/10.3390/genes14040777>
- Obanawa, H., Yoshitoshi, R., Watanabe, N., & Sakanoue, S. (2020). Portable LiDAR-Based Method for Improvement of Grass Height Measurement Accuracy: Comparison with SfM Methods. *Sensors*, 20(17), Artigo 17. <https://doi.org/10.3390/s20174809>
- Octariady, J., Hikmat, A., Widyaningrum, E., Mayasari, R., & Fajari, M. K. (2017). Vertical Accuracy Comparison of Digital Elevation Model from Lidar And Multitemporal Satellite Imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-1/W1, 419–423. <https://doi.org/10.5194/isprs-archives-XLII-1-W1-419-2017>

- Ornelas, A., Delgado-Vences, F., Morales-Bojórquez, E., Cruz-Escalona, V. H., Marín-Enríquez, E., & Hernández-Camacho, C. J. (2023). Modeling the biological growth with a random logistic differential equation. *Environmental and Ecological Statistics*, 30(2), 233–260. <https://doi.org/10.1007/s10651-023-00561-y>
- Orynbaikyzy, A., Gessner, U., & Conrad, C. (2022). Spatial Transferability of Random Forest Models for Crop Type Classification Using Sentinel-1 and Sentinel-2. *Remote Sensing*, 14(6), 1493. <https://doi.org/10.3390/rs14061493>
- Pagano, M. C., & Miransari, M. (2016). The importance of soybean production worldwide. Em *Abiotic and Biotic Stresses in Soybean Production* (p. 1–26). Elsevier. <https://doi.org/10.1016/B978-0-12-801536-0.00001-3>
- Pereira, B. M., & Brisola, M. V. (2022). Techno-economic Evolution of Soybean Production in Brazil and Argentina. *Journal of Agricultural Science*, 14(8), 145. <https://doi.org/10.5539/jas.v14n8p145>
- Pugh, N. A., Horne, D. W., Murray, S. C., Carvalho, G., Malambo, L., Jung, J., Chang, A., Maeda, M., Popescu, S., Chu, T., Starek, M. J., Brewer, M. J., Richardson, G., & Rooney, W. L. (2018). Temporal Estimates of Crop Growth in Sorghum and Maize Breeding Enabled by Unmanned Aerial Systems. *The Plant Phenome Journal*, 1(1), 1–10. <https://doi.org/10.2135/tppj2017.08.0006>
- R Core Team. (2023). *R: A Language and Environment for Statistical Computing* [Software]. <https://www.R-project.org/>
- Rabelo, W. P. S., Cruz, G., Silva, A. V. S., Rosa, F. C., & Gomes, W. C. (2023). Investigating the socioeconomic implications of the soybean production chain in Brazilian Northeast, Maranhão State: Investigando as implicações socioeconômicas da cadeia produtiva da soja no Nordeste brasileiro, Estado do Maranhão. *Concilium*, 23(12), 151–171. <https://doi.org/10.53660/CLM-1420-23H07>
- Radočaj, D., Šiljeg, A., Marinović, R., & Jurišić, M. (2023). 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>
- Ramezan, C. A., Warner, T. A., Maxwell, A. E., & Price, B. S. (2021). Effects of Training Set Size on Supervised Machine-Learning Land-Cover Classification of Large-Area High-Resolution Remotely Sensed Data. *Remote Sensing*, 13(3), 368. <https://doi.org/10.3390/rs13030368>

- Rawat, S. S., & Mishra, A. K. (2022). *Review of Methods for Handling Class-Imbalanced in Classification Problems* (Versão 1). arXiv. <https://doi.org/10.48550/ARXIV.2211.05456>
- Regier, P., Duggan, M., Myers-Pigg, A., & Ward, N. (2023). Effects of random forest modeling decisions on biogeochemical time series predictions. *Limnology and Oceanography: Methods*, 21(1), 40–52. <https://doi.org/10.1002/lom3.10523>
- Rowell, J. (2023). *Semantic Validation in Structure from Motion* (Versão 1). arXiv. <https://doi.org/10.48550/ARXIV.2304.02420>
- Salach, A., Bakula, K., Pilarska, M., Ostrowski, W., Górski, K., & Kurczyński, Z. (2018). Accuracy Assessment of Point Clouds from LiDAR and Dense Image Matching Acquired Using the UAV Platform for DTM Creation. *ISPRS International Journal of Geo-Information*, 7(9), 342. <https://doi.org/10.3390/ijgi7090342>
- Salokhe, M. R. (2023). Machine Learning: Applications in Agriculture (Crop Yield Prediction, Disease and Pest Detection). *International Journal of Advanced Research in Science, Communication and Technology*, 592–597. <https://doi.org/10.48175/IJAR SCT-12088>
- Sarkar, S., Zhou, J., Scaboo, A., Zhou, J., Aloysius, N., & Lim, T. T. (2023). Assessment of Soybean Lodging Using UAV Imagery and Machine Learning. *Plants*, 12(16), 2893. <https://doi.org/10.3390/plants12162893>
- Schulthess, U., Rodrigues, F., Taymans, M., Bellemans, N., Bontemps, S., Ortiz-Monasterio, I., Gérard, B., & Defourny, P. (2023). Optimal Sample Size and Composition for Crop Classification with Sen2-Agri's Random Forest Classifier. *Remote Sensing*, 15(3), 608. <https://doi.org/10.3390/rs15030608>
- Sediyama, T., Silva, F., & Borém, A. (2015). *Soja: Do plantio à colheita*. Viçosa: UFV.
- Senn, J. A., Mills, J., Walsh, C. L., Addy, S., & Peppas, M.-V. (2022, março 25). *Assessment of sensor pre-calibration to mitigate systematic errors in SfM photogrammetric surveys*. <https://doi.org/10.5194/egusphere-egu22-279>
- Shu, M., Fei, S., Zhang, B., Yang, X., Guo, Y., Li, B., & Ma, Y. (2022). Application of UAV Multisensor Data and Ensemble Approach for High-Throughput Estimation of Maize Phenotyping Traits. *Plant Phenomics*, 2022, 2022/9802585. <https://doi.org/10.34133/2022/9802585>
- Sun, M., Zhao, K., Wang, J., Mu, W., Zhan, Y., Li, W., Teng, W., Zhao, X., & Han, Y. (2022). Identification of quantitative trait loci underlying lodging of soybean

- across multiple environments. *Crop & Pasture Science*, 73(6), 652–662.
<https://doi.org/10.1071/CP21468>
- Tomohiro, K., & Homma, K. (2023). *Impact Assessment of Main Stem Elongation and Wind Speed on Lodging of Soybean Cultivar “Miyagishirome”*.
<https://doi.org/10.21203/rs.3.rs-3103937/v1>
- Triandita, N., & Putri, N. E. (2019). THE ROLE OF SOYBEAN IN CONTROL OF DEGENERATIVE DISEASE. *Jurnal Teknologi Pengolahan Pertanian*, 1(1), Artigo 1. <https://doi.org/10.35308/jtpp.v1i1.1478>
- Tulldahl, M., Rydell, J., Holmgren, J., Nordlöf, J., & Willén, E. (2019). Lidar-based positioning in forest environments. Em G. W. Kamerman & O. Steinvall (Orgs.), *Electro-Optical Remote Sensing XIII* (p. 4). SPIE.
<https://doi.org/10.1117/12.2533299>
- Volpato, L., Dobbels, A., Borem, A., & Lorenz, A. J. (2021). Optimization of temporal UAS-based imagery analysis to estimate plant maturity date for soybean breeding. *The Plant Phenome Journal*, 4(1), e20018. <https://doi.org/10.1002/ppj2.20018>
- Volpato, L., Pinto, F., González-Pérez, L., Thompson, I. G., Borém, A., Reynolds, M., Gérard, B., Molero, G., & Rodrigues, F. A. (2021). High Throughput Field Phenotyping for Plant Height Using UAV-Based RGB Imagery in Wheat Breeding Lines: Feasibility and Validation. *Frontiers in Plant Science*, 12, 591587. <https://doi.org/10.3389/fpls.2021.591587>
- Wang, R., Qiu, Y., Zhou, Y., Liang, Z., & Schnable, J. C. (2020). A High-Throughput Phenotyping Pipeline for Image Processing and Functional Growth Curve Analysis. *Plant Phenomics*, 2020, 2020/7481687. <https://doi.org/10.34133/2020/7481687>
- Wickham, H. (2016). Getting Started with ggplot2. Em H. Wickham (Org.), *Ggplot2: Elegant Graphics for Data Analysis* (p. 11–31). Springer International Publishing.
https://doi.org/10.1007/978-3-319-24277-4_2
- Wilke, N., Siegmann, B., Frimpong, F., Muller, O., Klingbeil, L., & Rascher, U. (2019). Quantifying Lodging Percentage, Lodging Development and Lodging Severity Using a Uav-Based Canopy Height Model. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4213, 649–655. <https://doi.org/10.5194/isprs-archives-XLII-2-W13-649-2019>

- Woebbecke, D. M., Meyer, G. E., Von Bargen, K., & Mortensen, D. A. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the American Society of Agricultural Engineers*, 38(1), 259–269.
- Wu, W., Shah, F., & Ma, B. (2022). Understanding of crop lodging and agronomic strategies to improve the resilience of rapeseed production to climate change. *Crop and Environment*, 1(2), 133–144. <https://doi.org/10.1016/j.crope.2022.05.005>
- Xu, Y., Zhang, R., Hou, Z., Yan, C., Xia, X., Ma, C., Dong, S., & Gong, Z. (2020). Mechanical properties of soybean plants under various plant densities. *Crop and Pasture Science*, 71(3), 249–259. <https://doi.org/10.1071/CP19133>
- Yang, M., Gao, P., Zhou, P., Xie, J., Sun, D., Han, X., & Wang, W. (2021). Simulating Canopy Temperature Using a Random Forest Model to Calculate the Crop Water Stress Index of Chinese Brassica. *Agronomy*, 11(11), 2244. <https://doi.org/10.3390/agronomy11112244>
- Yin, X. (2003). A Flexible Sigmoid Function of Determinate Growth. *Annals of Botany*, 91(3), 361–371. <https://doi.org/10.1093/aob/mcg029>
- Zheng, C., Guo, S., & Kern, J. L. (2021). Fast Bayesian Estimation for the Four-Parameter Logistic Model (4PLM). *SAGE Open*, 11(4), 215824402110525. <https://doi.org/10.1177/21582440211052556>

APÊNDICE

Apêndice 1 – Caracterização das alturas do dossel estimadas por análise de imagens (ALTDI) estimados em cada voo no ensaio 1 (E1). Valores médios e desvio-padrão calculados a partir de todos os valores dos percentis 50% (P50), 75% (P75), 90% (P90), 95% (P95) e 99% (P99). Intervalo modal e frequência da ocorrência do intervalo de altura predominante (IAP) dentre todas as parcelas.

Número do Voo (DAP)	P50 – Média (± Desvio-padrão)	P75 – Média (± Desvio-padrão)	P90 – Média (± Desvio-padrão)	P95 – Média (± Desvio-padrão)	P99 – Média (± Desvio-padrão)	IAP – Moda (Frequência)
1 (57)	80,2 (±12,0)	84,4 (±11,5)	87,4 (±11,2)	88,8 (±11)	90,7 (±10,9)	80 a 90 (0,29)
2 (64)	92,6 (±12,7)	98,4 (±11,8)	102,7 (±12,1)	104,9 (±12,5)	108,4 (±13,1)	90 a 100 (0,39)
3 (71)	89,8 (±14,1)	95,0 (±13,7)	98,5 (±13,5)	100,3 (±13,4)	102,7 (±13,5)	90 a 100 (0,26)
4 (78)	106,4 (±14,3)	112,6 (±13,2)	116,8 (±12,8)	118,8 (±12,6)	121,8 (±12,6)	100 a 110 (0,27)
5 (92)	106,8 (±16,1)	114,3 (±15,7)	119,2 (±15,6)	121,3 (±15,4)	124,1 (±15,0)	110 a 120 (0,22)
6 (100)	94,9 (±15,9)	102,8 (±15,6)	108,6 (±15,8)	111,2 (±16,2)	114,6 (±16,4)	90 a 100 (0,21)
7 (113)	81,5 (±17,5)	91,6 (±17,4)	100,0 (±17,5)	104,3 (±17,9)	110,3 (±19,2)	70 a 80 (0,24)
8 (120)	71,7 (±20,5)	81,0 (±20,9)	88,4 (±21,5)	91,8 (±21,8)	96,6 (±22,1)	60 a 70 (0,20)
9 (124)	82,3 (±16,9)	92,3 (±17,1)	100,4 (±17,4)	104,5 (±18,0)	110,4 (±19,2)	80 a 90 (0,24)
10 (127)	77,9 (±18,2)	88,3 (±18,2)	97,0 (±18,6)	101,5 (±19,1)	107,3 (±19,9)	70 a 80 (0,32)
11 (131)	74,2 (±17,6)	82,7 (±18,3)	89,7 (±18,8)	92,5 (±19,1)	96,2 (±19,6)	60 a 70 (0,23)
12 (135)	53,6 (±19,3)	63,7 (±20,9)	71,7 (±21,9)	76,0 (±22,6)	82,5 (±23,8)	50 a 60 (0,18)
13 (142)	46,8 (±19,2)	56,4 (±21,2)	64,2 (±22,7)	68,4 (±23,4)	73,8 (±24,8)	20 a 30 (0,21)
14 (145)	50,3 (±20,8)	60,4 (±23,2)	68,4 (±24,3)	72,6 (±24,9)	79,3 (±25,9)	40 a 50 (0,18)
15 (147)	39,2 (±18,3)	49,2 (±21,5)	57,1 (±23,3)	61,3 (±24,6)	67,7 (±25,3)	10 a 20 (0,21)
16 (153)	36,2 (±18,5)	45,0 (±20,7)	52,3 (±22,3)	56,0 (±22,9)	61,4 (±23,2)	20 a 30 (0,25)
17 (154)	35,0 (±15,7)	42,1 (±17,3)	47,1 (±18,0)	49,9 (±18,6)	53,5 (±18,7)	20 a 30 (0,24)
18 (156)	25,8 (±13,0)	32,1 (±15,4)	37,2 (±16,6)	39,7 (±16,9)	43,6 (±17,7)	10 a 20 (0,45)
19 (161)	28,8 (±13,6)	36,4 (±15,9)	43,0 (±17,8)	46,0 (±18,2)	49,6 (±18,8)	10 a 20 (0,32)

Apêndice 2 – Caracterização das alturas do dossel estimadas por análise de imagens (ALTDI) estimados em cada voo no ensaio 2 (E2). Valores médios e desvio-padrão calculados a partir de todos os valores dos percentis 50% (P50), 75% (P75), 90% (P90), 95% (P95) e 99% (P99). Intervalo modal e frequência da ocorrência do intervalo de altura predominante (IAP) dentre todas as parcelas.

Número do Voo (DAP)	P50 – Média (± Desvio-padrão)	P75 – Média (± Desvio-padrão)	P90 – Média (± Desvio-padrão)	P95 – Média (± Desvio-padrão)	P99 – Média (± Desvio-padrão)	IAP – Moda (Frequência)
1 (45)	74,0 (±10,6)	78,5 (±10,4)	81,5 (±10,4)	82,8 (±10,4)	84,5 (±10,5)	70 a 80 (0,36)
2 (52)	87,5 (±10,6)	92,6 (±10,6)	95,9 (±10,7)	97,5 (±10,8)	99,7 (±10,9)	90 a 100 (0,34)
3 (59)	93,0 (±11,8)	98,7 (±11,8)	103,2 (±11,9)	105,2 (±12,1)	108,3 (±12,4)	90 a 100 (0,31)
4 (66)	100,1 (±14,1)	107,0 (±13,1)	111,9 (±12,6)	114,0 (±12,6)	117,1 (±12,8)	100 a 110 (0,28)
5 (80)	102,9 (±15,9)	111,3 (±15,4)	117,2 (±15,0)	119,4 (±14,9)	122,1 (±14,7)	100 a 110 (0,22)
6 (88)	93,4 (±15,6)	102,3 (±16,0)	109,1 (±16,3)	111,9 (±16,3)	115,5 (±16,4)	90 a 100 (0,21)
7 (101)	81,7 (±17,1)	91,5 (±18,0)	99,9 (±18,7)	104,2 (±19,1)	110,0 (±19,9)	70 a 80 (0,21)
8 (108)	72,5 (±15,0)	82,0 (±16,1)	89,6 (±16,8)	93,0 (±17,3)	97,8 (±17,9)	60 a 70 (0,25)
9 (112)	82,5 (±16,4)	91,9 (±17,6)	100,2 (±18,5)	104,3 (±19,0)	110,1 (±19,9)	70 a 80 (0,21)
10 (115)	77,0 (±18,1)	87,3 (±19,4)	96,8 (±20,3)	101,1 (±20,9)	107,7 (±21,6)	70 a 80 (0,22)
11 (119)	73,2 (±16,9)	81,8 (±18,0)	88,8 (±19,2)	92,1 (±19,6)	96,4 (±20,5)	60 a 70 (0,23)
12 (123)	56,2 (±20,1)	66,2 (±21,4)	74,5 (±22,7)	78,5 (±23,5)	84,5 (±24,3)	50 a 60 (0,21)
13 (130)	41,8 (±16,8)	50,5 (±19,8)	57,8 (±21,5)	62,1 (±22,1)	67,7 (±23,1)	20 a 30 (0,28)
14 (133)	38,7 (±17,4)	49,0 (±19,8)	58,5 (±22,8)	63,1 (±24,2)	69,6 (±25,3)	20 a 30 (0,23)
15 (135)	31,5 (±13,8)	40,3 (±17,8)	47,8 (±19,9)	51,7 (±20,8)	56,9 (±21,7)	20 a 30 (0,34)
16 (141)	33,6 (±14,2)	40,8 (±17,2)	47,9 (±19,1)	51,3 (±19,4)	56,0 (±19,7)	20 a 30 (0,32)
17 (142)	32,4 (±14,0)	39,3 (±16,7)	44,1 (±18,0)	46,4 (±18,3)	49,6 (±18,6)	20 a 30 (0,30)
18 (144)	18,0 (±8,9)	23,1 (±11,4)	27,6 (±13,2)	29,7 (±13,8)	32,8 (±14,4)	10 a 20 (0,69)
19 (149)	26,5 (±10,4)	32,0 (±12,1)	36,4 (±13,2)	38,7 (±13,6)	42,2 (±14,0)	20 a 30 (0,36)

Apêndice 3 – Estatísticas descritivas dos parâmetros gerados pelo ajuste da curva logística às alturas do dossel normalizadas pelo valor máximo (N_ALTDI) nos diferentes níveis de acamamento (AC), estimados segundo os métodos de estimação da altura do dossel por análise de imagens (ME_ALTDI) dos percentis 50% (P50), 75% (P75), 90% (P90), 95% (P95), 99% (P99) e intervalo de altura predominante (IAP).

ME_ALTDI	Parâmetro	Estatística	AC				
			1	2	3	4	5
P50	L	3º Quartil	0,86	0,82	0,81	0,78	0,72
		Mediana	0,64	0,50	0,56	0,58	0,57
		1º Quartil	0,31	0,30	0,31	0,33	0,44
	k	3º Quartil	-0,17	-0,18	-0,17	-0,13	-0,15
		Mediana	-0,52	-0,60	-0,28	-0,23	-0,28
		1º Quartil	-3,11	-2,85	-1,02	-1,72	-1,83
	x0	3º Quartil	-6,04	-7,50	-6,49	-5,53	-11,41
		Mediana	-12,26	-13,78	-14,09	-13,76	-19,56
		1º Quartil	-20,27	-19,50	-21,54	-20,56	-24,53
	b	3º Quartil	0,62	0,64	0,58	0,53	0,46
		Mediana	0,30	0,43	0,37	0,31	0,30
		1º Quartil	0,04	0,09	0,00	0,00	0,10
	R2	3º Quartil	0,82	0,82	0,86	0,87	0,84
		Mediana	0,70	0,68	0,74	0,78	0,78
		1º Quartil	0,47	0,53	0,58	0,60	0,61
P75	L	3º Quartil	0,91	0,82	0,78	0,72	0,64
		Mediana	0,50	0,45	0,50	0,47	0,48
		1º Quartil	0,28	0,22	0,30	0,29	0,33
	k	3º Quartil	-0,17	-0,19	-0,16	-0,14	-0,16
		Mediana	-0,43	-0,69	-0,31	-0,29	-0,32
		1º Quartil	-2,13	-2,86	-1,79	-2,60	-3,35
	x0	3º Quartil	-6,98	-8,04	-6,47	-7,93	-11,33
		Mediana	-12,08	-14,56	-13,88	-14,95	-19,16
		1º Quartil	-22,29	-20,12	-23,62	-22,07	-25,64
	b	3º Quartil	0,69	0,75	0,65	0,66	0,55
		Mediana	0,50	0,54	0,50	0,49	0,43
		1º Quartil	0,13	0,20	0,16	0,17	0,30
	R2	3º Quartil	0,78	0,78	0,82	0,86	0,83
		Mediana	0,66	0,62	0,71	0,76	0,75
		1º Quartil	0,39	0,45	0,49	0,51	0,55

Continuação

ME_ALTDI	Parâmetro	Estatística	AC				
			1	2	3	4	5
P90	L	3º Quartil	0,86	0,82	0,93	0,68	0,52
		Mediana	0,44	0,39	0,48	0,44	0,40
		1º Quartil	0,26	0,23	0,30	0,28	0,27
	k	3º Quartil	-0,16	-0,20	-0,15	-0,15	-0,22
		Mediana	-0,61	-0,64	-0,38	-0,31	-0,49
		1º Quartil	-3,27	-2,91	-2,07	-1,81	-2,63
	x0	3º Quartil	-6,43	-6,67	-6,24	-7,69	-13,18
		Mediana	-11,98	-14,24	-13,65	-15,27	-19,35
		1º Quartil	-22,22	-21,02	-23,56	-23,73	-23,26
	b	3º Quartil	0,76	0,79	0,70	0,71	0,68
		Mediana	0,60	0,64	0,54	0,55	0,57
		1º Quartil	0,21	0,27	0,16	0,32	0,43
	R2	3º Quartil	0,74	0,73	0,76	0,83	0,80
		Mediana	0,62	0,57	0,63	0,70	0,68
		1º Quartil	0,35	0,34	0,41	0,36	0,54
P95	L	3º Quartil	0,96	0,87	0,91	0,71	0,56
		Mediana	0,46	0,40	0,44	0,45	0,43
		1º Quartil	0,29	0,24	0,27	0,29	0,30
	k	3º Quartil	-0,10	-0,18	-0,15	-0,16	-0,23
		Mediana	-0,50	-0,58	-0,39	-0,34	-0,39
		1º Quartil	-3,14	-2,14	-2,05	-2,46	-1,74
	x0	3º Quartil	-6,01	-7,57	-5,78	-8,57	-14,30
		Mediana	-11,36	-14,50	-13,24	-15,84	-19,35
		1º Quartil	-23,31	-24,13	-23,67	-25,17	-26,47
	b	3º Quartil	0,76	0,82	0,74	0,74	0,70
		Mediana	0,61	0,67	0,61	0,58	0,56
		1º Quartil	0,16	0,26	0,23	0,30	0,40
	R2	3º Quartil	0,72	0,72	0,72	0,80	0,80
		Mediana	0,56	0,53	0,60	0,68	0,65
		1º Quartil	0,28	0,29	0,37	0,39	0,50

Continuação

ME_ALTDI	Parâmetro	Estatística	AC				
			1	2	3	4	5
P99	L	3º Quartil	0,91	0,77	0,91	0,72	0,55
		Mediana	0,42	0,36	0,42	0,43	0,45
		1º Quartil	0,28	0,20	0,27	0,25	0,26
	k	3º Quartil	0,09	-0,15	-0,16	-0,13	-0,11
		Mediana	-0,41	-0,63	-0,40	-0,38	-0,32
		1º Quartil	-2,20	-2,93	-1,97	-3,02	-1,54
	x0	3º Quartil	-7,13	-8,00	-8,13	-9,15	-15,19
		Mediana	-12,23	-15,21	-16,00	-16,60	-19,53
		1º Quartil	-58,10	-27,94	-28,64	-28,66	-30,49
	b	3º Quartil	0,83	0,90	0,78	0,79	0,76
		Mediana	0,66	0,74	0,67	0,64	0,60
		1º Quartil	0,35	0,36	0,28	0,39	0,47
	R2	3º Quartil	0,70	0,70	0,70	0,81	0,77
		Mediana	0,57	0,48	0,54	0,64	0,64
		1º Quartil	0,31	0,24	0,32	0,37	0,47
IAP	L	3º Quartil	0,78	0,76	0,80	0,82	0,80
		Mediana	0,59	0,42	0,49	0,56	0,61
		1º Quartil	0,27	0,25	0,26	0,29	0,36
	k	3º Quartil	-0,18	-0,18	-0,15	-0,12	-0,13
		Mediana	-0,82	-0,85	-0,38	-0,31	-0,26
		1º Quartil	-3,88	-3,58	-2,27	-2,12	-2,66
	x0	3º Quartil	-7,99	-9,03	-7,92	-8,23	-10,62
		Mediana	-13,19	-14,51	-14,63	-15,03	-18,14
		1º Quartil	-21,59	-21,72	-22,33	-23,23	-26,98
	b	3º Quartil	0,56	0,58	0,57	0,55	0,52
		Mediana	0,19	0,41	0,36	0,29	0,20
		1º Quartil	0,00	0,00	0,00	0,00	0,06
	R2	3º Quartil	0,80	0,81	0,82	0,85	0,85
		Mediana	0,66	0,66	0,73	0,76	0,74
		1º Quartil	0,40	0,45	0,51	0,53	0,60

Apêndice 4 – Médias dos valores de alturas do dossel normalizadas pelo valor máximo (N_ALTDI) das séries temporais intervalos de 7 dias relativos da execução do voo (DIASV), entre os DIASV -105 e 0, obtidos dos métodos de estimação da altura do dossel por análise de imagens (ME_ALTDI) dos percentis 50% (P50), 75% (P75), 90% (P90), 95% (P95), 99% (P99) e intervalo de altura predominante (IAP).

ME_ALTDI	Intervalo de DIASV	AC				
		1	2	3	4	5
P50	-98 a -105	0,6274	0,5932	0,7208	0,6521	0,6421
	-91 a -98	0,6406	0,6041	0,7308	0,7210	0,6974
	-84 a -91	0,6579	0,6759	0,7428	0,7531	0,7410
	-77 a -84	0,7002	0,7412	0,7560	0,7986	0,8126
	-70 a -77	0,7356	0,7821	0,8091	0,8259	0,8489
	-63 a -70	0,7756	0,8301	0,8373	0,8549	0,8121
	-56 a -63	0,8291	0,8717	0,8743	0,8865	0,8444
	-49 a -56	0,8774	0,8841	0,9036	0,8769	0,8331
	-42 a -49	0,8876	0,8837	0,8732	0,8621	0,8677
	-35 a -42	0,8993	0,8886	0,8735	0,8290	0,8077
	-28 a -35	0,8627	0,8427	0,8414	0,7791	0,7271
	-21 a -28	0,8207	0,7565	0,7546	0,7620	0,6359
	-14 a -21	0,7317	0,6802	0,6587	0,6523	0,5476
	-7 a -14	0,6233	0,6046	0,6080	0,5588	0,4789
	0 a -7	0,4512	0,4766	0,4625	0,4460	0,3569
P75	-98 a -105	0,6245	0,5709	0,6987	0,6387	0,6626
	-91 a -98	0,6371	0,5928	0,7125	0,7082	0,6995
	-84 a -91	0,6627	0,6633	0,7297	0,7393	0,7322
	-77 a -84	0,7018	0,7283	0,7430	0,7821	0,8005
	-70 a -77	0,7353	0,7680	0,7940	0,8130	0,8398
	-63 a -70	0,7770	0,8198	0,8259	0,8496	0,8273
	-56 a -63	0,8265	0,8682	0,8771	0,8892	0,8606
	-49 a -56	0,8845	0,8921	0,9059	0,8889	0,8410
	-42 a -49	0,8923	0,9043	0,8953	0,8944	0,8569
	-35 a -42	0,9097	0,9065	0,8915	0,8483	0,8229
	-28 a -35	0,8788	0,8668	0,8620	0,8105	0,7545
	-21 a -28	0,8514	0,7891	0,7740	0,7866	0,6759
	-14 a -21	0,7688	0,7198	0,6911	0,6794	0,5856
	-7 a -14	0,6681	0,6462	0,6510	0,6095	0,5333
	0 a -7	0,5177	0,5527	0,5245	0,5102	0,4340

Continuação

ME_ALTDI	Intervalo de DIASV	AC				
		1	2	3	4	5
P90	-98 a -105	0,6214	0,5537	0,6768	0,6222	0,6783
	-91 a -98	0,6262	0,5803	0,6815	0,6947	0,7094
	-84 a -91	0,6589	0,6557	0,7145	0,7245	0,7345
	-77 a -84	0,6966	0,7197	0,7293	0,7653	0,7935
	-70 a -77	0,7319	0,7563	0,7792	0,8010	0,8232
	-63 a -70	0,7743	0,8126	0,8163	0,8412	0,8291
	-56 a -63	0,8251	0,8518	0,8672	0,8868	0,8698
	-49 a -56	0,8848	0,8879	0,9044	0,9066	0,8624
	-42 a -49	0,8965	0,9110	0,9017	0,9080	0,8481
	-35 a -42	0,9171	0,9120	0,8954	0,8654	0,8244
	-28 a -35	0,8964	0,8725	0,8686	0,8277	0,8002
	-21 a -28	0,8702	0,8039	0,7931	0,7988	0,7096
	-14 a -21	0,7981	0,7490	0,7163	0,7011	0,6264
	-7 a -14	0,7095	0,6798	0,6845	0,6534	0,5798
	0 a -7	0,5709	0,5986	0,5705	0,5603	0,4911
P95	-98 a -105	0,6240	0,5472	0,6519	0,6160	0,6718
	-91 a -98	0,6230	0,5719	0,6685	0,6873	0,7020
	-84 a -91	0,6601	0,6501	0,7066	0,7188	0,7306
	-77 a -84	0,6944	0,7140	0,7238	0,7596	0,7854
	-70 a -77	0,7307	0,7507	0,7712	0,7966	0,8179
	-63 a -70	0,7715	0,8096	0,8095	0,8389	0,8243
	-56 a -63	0,8240	0,8462	0,8647	0,8823	0,8668
	-49 a -56	0,8871	0,8835	0,9007	0,9071	0,8784
	-42 a -49	0,8968	0,9105	0,9031	0,9130	0,8573
	-35 a -42	0,9186	0,9141	0,8985	0,8740	0,8345
	-28 a -35	0,9040	0,8749	0,8733	0,8406	0,8096
	-21 a -28	0,8813	0,8099	0,8068	0,8071	0,7190
	-14 a -21	0,8129	0,7592	0,7289	0,7176	0,6356
	-7 a -14	0,7270	0,6957	0,6986	0,6709	0,5977
	0 a -7	0,5981	0,6215	0,5934	0,5815	0,5187

Continuação

ME_ALTDI	Intervalo de DIASV	AC				
		1	2	3	4	5
P99	-98 a -105	0,6149	0,5419	0,6278	0,6041	0,6668
	-91 a -98	0,6111	0,5705	0,6575	0,6803	0,6840
	-84 a -91	0,6567	0,6471	0,6978	0,7112	0,7115
	-77 a -84	0,6883	0,7097	0,7134	0,7495	0,7748
	-70 a -77	0,7266	0,7446	0,7595	0,7854	0,8013
	-63 a -70	0,7670	0,8047	0,7982	0,8279	0,8126
	-56 a -63	0,8210	0,8428	0,8584	0,8753	0,8647
	-49 a -56	0,8869	0,8829	0,8961	0,9094	0,8921
	-42 a -49	0,8989	0,9122	0,9036	0,9158	0,8701
	-35 a -42	0,9187	0,9181	0,8989	0,8755	0,8416
	-28 a -35	0,9080	0,8846	0,8767	0,8528	0,8077
	-21 a -28	0,8875	0,8197	0,8200	0,8239	0,7189
	-14 a -21	0,8257	0,7757	0,7427	0,7382	0,6462
	-7 a -14	0,7487	0,7181	0,7138	0,6897	0,6181
	0 a -7	0,6342	0,6585	0,6228	0,6029	0,5365
IAP	-98 a -105	0,6290	0,5907	0,6981	0,6679	0,6540
	-91 a -98	0,6519	0,6011	0,7440	0,7215	0,7009
	-84 a -91	0,6477	0,6815	0,7527	0,7517	0,7420
	-77 a -84	0,6995	0,7394	0,7647	0,7866	0,8001
	-70 a -77	0,7341	0,7866	0,8044	0,8240	0,8524
	-63 a -70	0,7785	0,8261	0,8270	0,8525	0,7883
	-56 a -63	0,8211	0,8632	0,8568	0,8770	0,8270
	-49 a -56	0,8737	0,8721	0,8985	0,8585	0,8755
	-42 a -49	0,8808	0,8842	0,8662	0,8304	0,8251
	-35 a -42	0,8927	0,8705	0,8604	0,8099	0,8058
	-28 a -35	0,8500	0,8446	0,8328	0,7680	0,7088
	-21 a -28	0,8168	0,7578	0,7478	0,7577	0,6367
	-14 a -21	0,7161	0,6741	0,6600	0,6531	0,5424
	-7 a -14	0,6048	0,6019	0,6112	0,5382	0,4801
	0 a -7	0,4432	0,4807	0,4723	0,4392	0,3144

Apêndice 5 – Estatísticas de desempenho por classe dos modelos de Random Forest (MRF) treinados com os dados dos parâmetros das curvas ajustadas à série temporal de alturas do dossel normalizadas pelo valor máximo (N_ALTDI), com conjuntos de treinamento dos ensaios 1 (E1) e 2 (E2), dos diferentes níveis de acamamento (AC).

MRF	AC	Treinamento com E1 e teste com E2				Treinamento com E2 e teste com E1			
		Acurácia Balanceada	F1-Score	Precisão	Sensibilidade	Acurácia Balanceada	F1-Score	Precisão	Sensibilidade
MRF_P50	1	0,6082	0,3974	0,3523	0,4559	0,6085	0,3455	0,2346	0,6552
	2	0,5100	0,2468	0,3519	0,1900	0,5195	0,1593	0,3600	0,1023
	3	0,5417	0,2645	0,3478	0,2133	0,5056	0,1406	0,3214	0,0900
	4	0,5684	0,2857	0,2464	0,3400	0,5466	0,3280	0,2793	0,3974
	5	0,5771	0,1290	0,0816	0,3077	0,5078	0,0625	0,0667	0,0588
MRF_P75	1	0,5599	0,3354	0,2903	0,3971	0,6007	0,3378	0,2778	0,4310
	2	0,5003	0,2405	0,3276	0,1900	0,4797	0,2613	0,2342	0,2955
	3	0,4471	0,1094	0,1321	0,0933	0,4999	0,2152	0,2931	0,1700
	4	0,5104	0,1964	0,1774	0,2200	0,5415	0,2838	0,3000	0,2692
	5	0,5522	0,1132	0,0750	0,2308	0,4815	0,0000	0,0000	0,0000
MRF_P90	1	0,5336	0,2435	0,2979	0,2059	0,5572	0,2970	0,2083	0,5172
	2	0,5254	0,3814	0,3565	0,4100	0,5010	0,1791	0,2609	0,1364
	3	0,4883	0,2469	0,2299	0,2667	0,5032	0,2092	0,3019	0,1600
	4	0,5480	0,2476	0,2364	0,2600	0,5455	0,3019	0,2963	0,3077
	5	0,4966	0,0000	0,0000	0,0000	0,5357	0,1176	0,1176	0,1176
MRF_P95	1	0,5347	0,2521	0,2941	0,2206	0,5171	0,2375	0,1863	0,3276
	2	0,5292	0,3542	0,3696	0,3400	0,4773	0,2692	0,2333	0,3182
	3	0,4894	0,2044	0,2258	0,1867	0,4667	0,1954	0,2297	0,1700
	4	0,5528	0,2677	0,2208	0,3400	0,5479	0,2456	0,3889	0,1795
	5	0,5795	0,1622	0,1250	0,2308	0,5480	0,1538	0,2222	0,1176
MRF_P99	1	0,5252	0,2182	0,2857	0,1765	0,5143	0,2222	0,1860	0,2759
	2	0,5365	0,3598	0,3820	0,3400	0,4911	0,2353	0,2439	0,2273
	3	0,4847	0,2667	0,2286	0,3200	0,5432	0,2981	0,3934	0,2400
	4	0,4884	0,1622	0,1475	0,1800	0,5528	0,3277	0,2929	0,3718
	5	0,5650	0,1818	0,2222	0,1538	0,5109	0,0667	0,0769	0,0588
MRF_IAP	1	0,4958	0,1935	0,2143	0,1765	0,5251	0,2570	0,1901	0,3966
	2	0,5208	0,3041	0,3662	0,2600	0,4901	0,2065	0,2388	0,1818
	3	0,4623	0,2299	0,2020	0,2667	0,4776	0,2235	0,2532	0,2000
	4	0,5026	0,1897	0,1667	0,2200	0,4807	0,1793	0,1940	0,1667
	5	0,5163	0,0741	0,0714	0,0769	0,5202	0,0833	0,1429	0,0588

MRF_P50: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 50%; MRF_P75: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 75%; MRF_P90: modelo treinado a partir das alturas do dossel estimadas por análise de imagens

como o percentil 90%; MRF_P95: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 95%; MRF_P99: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o percentil 99%; MRF_IAP: modelo treinado a partir das alturas do dossel estimadas por análise de imagens como o intervalo de altura predominante.