



JOÃO VICTOR REIS E SILVA

**RELATÓRIO DE ESTÁGIO: AVALIAÇÃO DA QUALIDADE
DE MUDAS PARA O PLANTIO MECANIZADO**

**LAVRAS – MG
2023**

JOÃO VICTOR REIS E SILVA

**RELATÓRIO DE ESTÁGIO: AVALIAÇÃO DA QUALIDADE DE MUDAS PARA O
PLANTIO MECANIZADO**

Relatório de estágio supervisionado apresentado à Universidade Federal de Lavras, como parte das exigências do Curso de Engenharia Florestal, para a obtenção do título de Bacharel.

Prof. Dr. Otávio Camargo Campoe
Orientador

**LAVRAS – MG
2023**

JOÃO VICTOR REIS E SILVA

**RELATÓRIO DE ESTÁGIO: AVALIAÇÃO DA QUALIDADE DE MUDAS PARA O
PLANTIO MECANIZADO**

**INTERNSHIP REPORT: EVALUATION OF THE QUALITY OF SEEDLINGS FOR THE
MECHANIZED PLANTING**

Relatório de estágio supervisionado apresentado à Universidade Federal de Lavras, como parte das exigências do Curso de Engenharia Florestal, para a obtenção do título de Bacharel.

APROVADO em 24 de fevereiro de 2023

Prof. Dr. Otávio Camargo Campoe, UFLA

Márcio de Souza Reis, CENIBRA

Geraldo Magela Pereira Filho, CENIBRA

Prof. Dr. Otávio Camargo Campoe

Orientador

LAVRAS – MG

2023

Ao meu pai, Júlio, que me ensinou tanta coisa.

À minha mãe, Elisabete, que se doou tanto para que eu realizasse meu sonho.

Às minhas tias, Estelita, Irene e Iris, que cuidaram de mim com muito amor quando meus pais mais precisaram.

Ao meu irmão caçula, Gustavo, por não me deixar desistir.

Dedico

AGRADECIMENTOS

À Deus pelo dom da vida.

Aos meus pais, Júlio e Elisabete, e irmão, Gustavo, por serem minha fonte de força e inspiração.

À toda minha família pelo apoio incondicional.

À Universidade Federal de Lavras por me proporcionar a oportunidade de realizar o sonho da graduação.

A todos os professores da universidade por compartilharem conhecimento de uma maneira tão didática e esclarecedora. Em especial aos professores do Departamento de Ciências Florestais.

Ao professor Dr. Otávio Camargo Campoe pela orientação e incentivo na execução do trabalho.

Aos meus amigos Allan, Eduardo, Letícia, Luiza, Luiz e Nicolas por todos os momentos de descontração, conversas e estudo.

À Celulose Nipo-Brasileira S.A. – CENIBRA pela oportunidade de estágio.

A toda equipe do Departamento de Silvicultura da CENIBRA pela receptividade e ensinamentos compartilhados. Em especial ao meu supervisor de estágio, Henrique, por todo o incentivo, paciência e liberdade que foram fundamentais para o meu desenvolvimento.

A todos, o meu muito obrigado!!!

RESUMO

O presente trabalho teve como objetivo realizar a avaliação de parâmetros morfológicos das mudas de dois materiais genéticos comerciais da empresa Celulose Nipo-Brasileira S.A. (CENIBRA), com o intuito de identificar qual parâmetro impacta negativamente na qualidade da operação de plantio mecanizado praticado na empresa. Os parâmetros morfológicos avaliados foram altura da muda, diâmetro do colo, número de par de folhas expandidas, existência de bifurcação, agregação do substrato, posicionamento da muda no tubete e inclinação basal. As informações foram coletadas e registradas através do aplicativo mobile ArcGis Survey123 e, posteriormente, analisadas com o auxílio do *software* RStudio por meio do método de regressão logística. Três modelos de regressão foram gerados: generalizado; para o Clone A; e para o Clone B. Os principais parâmetros morfológicos que influenciaram negativamente na qualidade da atividade foram o diâmetro do colo, inclinação basal e o posicionamento da muda no tubete. Para o Clone A, o diâmetro e a inclinação foram os principais enquanto que, para o Clone B, a inclinação basal e o posicionamento da muda influenciaram mais. Os modelos ajustados apresentaram boa performance visto que a área sob a Curva ROC (Característica de Operação do Receptor) foi de 0,85 para o modelo generalizado; 0,87 para o modelo do Clone A; e 0,80 para o modelo do Clone B. Foi verificado que quanto maior é a inclinação basal das mudas, maior é a chance de novas falhas de plantio acontecerem; que as falhas de plantio têm maior probabilidade de ocorrência por mudas descentralizadas no tubete; e que quanto maior for o diâmetro do coleto, maior é a probabilidade de novas falhas surgirem.

Palavras-chave: Qualidade de muda. Falha de plantio. Plantio mecanizado.

ABSTRACT

The present work aimed to evaluate the morphological parameters of the seedlings of two commercial genetic materials from the company Celulose Nipo-Brasileira S.A. (CENIBRA), in order to identify which parameter has a negative impact on the quality of the mechanized planting operation practiced in the company. The morphological parameters evaluated were seedling height, collar diameter, number of expanded leaf pairs, existence of bifurcation, substrate aggregation, seedling position in the tube and basal inclination. The information was collected and recorded using the mobile application ArcGis Survey123 and subsequently analyzed using the RStudio software using the logistic regression method. Three regression models were generated: generalized; for Clone A; and for Clone B. The main morphological parameters that negatively influenced the quality of the activity were the collar diameter, basal inclination and the position of the seedling in the tube. For Clone A, the diameter and inclination were the main ones, while for Clone B, the basal inclination and the position of the seedling influenced more. The adjusted models showed good performance since the area under the ROC Curve (Receiver Operating Characteristic) was 0.85 for the generalized model; 0.87 for the Clone A model; and 0.80 for the Clone B model. It was found that the greater the basal inclination of the seedlings, the greater the chance of new planting failures occurring; that planting failures are more likely to occur due to decentralized seedlings in the tube; and that the greater the diameter of the neck, the greater the probability of new failures appearing.

Keywords: Seedling quality. Planting failure. Mechanized planting.

LISTA DE FIGURAS

Figura 1 – Exemplo de falhas de plantio: cova sem muda (a); cova com mais de uma muda (b)	11
Figura 2 – Estruturas do cabeçote: magazine (a); tubo de plantio (b).....	16
Figura 3 – Parâmetros morfológicos avaliados: altura (a, b); diâmetro do colo (c); inclinação basal (d); agregação do substrato (e, f); posicionamento da muda no tubete (g, h); bifurcação (i); número de par de folhas (j); avaliação do evento de interesse (k).....	18
Figura 4 – Interface do aplicativo ArcGis Survey123.....	19
Figura 5 – Efeitos de: DIAM (a); ANG (b); POS (c).....	24
Figura 6 – Curva ROC para o modelo generalizado.	25
Figura 7 – Efeitos de DIAM (a) e ANG (b) para o Clone A.	27
Figura 8 – Curva ROC para o modelo do Clone A.	28
Figura 9 – Efeitos de POS (a) e ANG (b) para o Clone B.....	29
Figura 10 – Curva ROC para o modelo do Clone B.....	30

LISTA DE TABELAS

Tabela 1 – Resumo dos dados	21
Tabela 2 – Estatísticas de ajuste do primeiro modelo generalizado.....	22
Tabela 3 – Estatísticas de ajuste do novo modelo generalizado.....	22
Tabela 4 – Razões de chance para o segundo modelo generalizado.	23
Tabela 5 – Matriz de classificação para o modelo generalizado.	25
Tabela 6 – Estatísticas de ajuste do modelo do Clone A.....	26
Tabela 7 – Razões de chance para o modelo do Clone A.....	26
Tabela 8 - Matriz de classificação para o modelo do Clone A.....	27
Tabela 9 – Estatísticas de ajuste do modelo do Clone B.....	28
Tabela 10 – Razões de chance para o modelo do Clone B.....	29
Tabela 11 – Matriz de classificação para o Clone B.	30

SUMÁRIO

1	INTRODUÇÃO.....	10
2	REFERENCIAL TEÓRICO.....	12
2.1	Plantio mecanizado	12
2.2	Regressão logística.....	13
3	MATERIAIS E MÉTODOS	15
3.1	A empresa	15
3.2	Mensuração dos parâmetros morfológicos das mudas.....	15
3.3	Registro e análise dos dados	19
4	RESULTADOS E DISCUSSÃO	21
4.1	Análise generalizada dos dados	21
4.1.1	Ajuste do modelo	21
4.1.2	Validação do modelo	24
4.1.2.1	Matriz de classificação.....	24
4.1.2.2	Curva ROC.....	25
4.2	Análise individualizada dos dados.....	26
5	CONSIDERAÇÕES FINAIS	31
6	CONCLUSÃO	32
	REFERÊNCIAS BIBLIOGRÁFICAS	34

1 INTRODUÇÃO

É sabido que o setor florestal brasileiro é um dos mais desenvolvidos no mundo. Esse desenvolvimento vem desde os anos 1910 quando as primeiras árvores de eucalipto foram plantadas para suprir as demandas da Companhia Paulista de Estradas de Ferro. A partir da década de 60, a área de florestas plantadas expandiu significativamente devido a alguns fatores como incentivos fiscais, por exemplo, que foram criados pelo governo do país a fim de reduzir os impactos antrópicos sobre a mata nativa. De acordo com Antonangelo e Bacha (1998) a área plantada no Brasil em 1964 era de aproximadamente 500 mil hectares. Atualmente, são pouco menos de 10 milhões de hectares (IBÁ, 2022).

De acordo com a Indústria Brasileira de Árvores (2022), o setor florestal apresentou uma receita bruta recorde no ano de 2021 com 244,6 bilhões de reais. Em relação a 2020, isso representou um crescimento de 7,5% que contribuiu para que o Produto Interno Bruto brasileiro crescesse 4,6% no mesmo ano. Esses números mostram como o setor é importante para a economia do país.

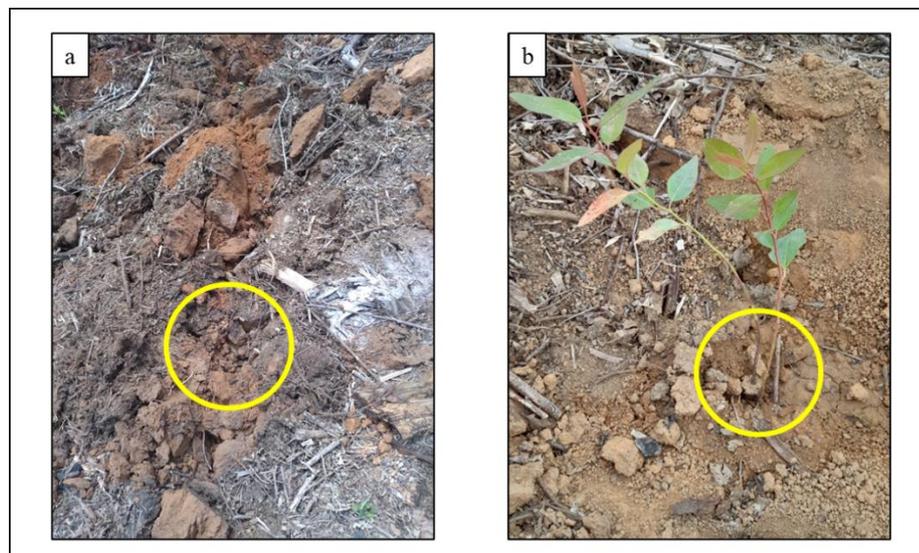
Nos dias atuais, as empresas têm procurado mecanizar o máximo possível de suas atividades com o intuito de obter maior rendimento e diminuir os custos com mão de obra. O Levantamento do Nível de Mecanização na Silvicultura realizado pelo Instituto de Pesquisas e Estudos Florestais (IPEF) (2021), classifica as empresas que cultivam eucalipto em 8 níveis (1 ao 8) de mecanização onde, o nível 1 indica que as atividades são realizadas de forma manual com a utilização de ferramentas sem acionamento mecânico e o nível 8 representa as atividades que são realizadas de forma autônoma, sem a necessidade de um operador em campo. De forma geral, considerando todas as atividades, as empresas se encontram no nível 5 de mecanização (trata-se de equipamentos que demandam de um operador e não utilizam piloto automático) com uma média de 42,08% de área trabalhada em tal nível no ano de 2020. Ao falar apenas da atividade de plantio, as empresas encontram-se no nível 2 (atividade manual com a utilização de ferramentas que possuem acionamento mecânico) de mecanização, com média de 56,88% de área trabalhada neste nível em 2020.

Para a Celulose Nipo-Brasileira (CENIBRA), empresa onde o presente trabalho foi desenvolvido, não é diferente. Devido às condições de relevo das áreas onde a empresa planta, o plantio é uma atividade que é feita, em sua maior parte, de maneira manual. No levantamento do nível de mecanização na silvicultura do IPEF, a CENIBRA está locada no nível 2 de mecanização quando o assunto é plantio.

Em meados de 2017 representantes da empresa estiveram presentes na Elmia Wood, a maior feira florestal do mundo, que foi sediada na Suécia e contou com diversos expositores de diversos países. No evento, eles conheceram um cabeçote de plantio mecanizado da fabricante sueca Bracke Forest, e vislumbraram a oportunidade de o trazerem para o Brasil. A empresa realizou a aquisição de uma unidade e desde então houve um trabalho árduo junto à fabricante para que o equipamento pudesse ser desenvolvido. Em 2022 a CENIBRA realizou a aquisição de mais cinco unidades do equipamento com o intuito de implementar um módulo de plantio mecanizado.

Durante a operação foram observadas algumas falhas na atividade, como por exemplo a ausência de mudas em determinadas covas e, em outras, mais de uma muda plantada (tais situações são ilustradas na Figura 1a e Figura 1b, respectivamente). Então, devido a essas falhas, viu-se a necessidade de realizar uma avaliação de impacto da qualidade das mudas que são destinadas ao módulo de plantio mecanizado, com o objetivo inicial de verificar quais parâmetros morfológicos das plantas poderiam ser mais restritivos para o sucesso da atividade e, a partir disso, definir um padrão de qualidade de mudas a ser indicado ao módulo de plantio mecanizado da empresa.

Figura 1 – Exemplo de falhas de plantio: cova sem muda (a); cova com mais de uma muda (b)



Fonte: Do autor (2022).

2 REFERENCIAL TEÓRICO

2.1 Plantio mecanizado

Manner & Ersson (2021) contam que o plantio mecanizado consiste em um equipamento que realiza o preparo do solo e depois realiza o plantio da muda. Os autores relatam ainda que durante a década de 1980, na Suécia, uma plantadora mecanizada realizou o plantio referente a até 12% de toda a área plantada nas regiões norte e central do país, chegando a plantar mais de 2.000 mudas por hora.

Na Finlândia, três plantadoras são utilizadas: Bracke; M-Planter; e Risutec. Elas são acopladas à lança de uma escavadora hidráulica, o operador determina os pontos de plantio e as mudas são depositadas em magazines que são abastecidos manualmente conforme necessidade (LAINE *et. al.*, 2016). Segundo os autores, o plantio mecanizado pode reduzir em até 22% a necessidade de recursos humanos.

De acordo com Soler *et. al.* (2019), no Brasil as formas mais comuns de realizar a operação de plantio é de maneira manual ou semimecanizada. Porém, os autores citam que uma alternativa a isso seria importar equipamentos já consolidados em outros países e adaptá-los às condições do Brasil.

Souza *et. al.* (2022) relatam que a mecanização contribuiria positivamente para elevar o patamar da operação de plantio ao levantarem os danos que a atividade manual causa na saúde dos trabalhadores, tais como: posturas inadequadas; elevado esforço físico; e existência do risco de lesão nas articulações.

Sereghetti (2016) afirma que o plantio mecanizado é sensível à qualidade das mudas e que mudas maiores podem causar entupimento do cabeçote de plantio, o que obriga o operador do equipamento a realizar a operação mais uma vez. O autor ainda relata que, para o plantio mecanizado, as mudas não podem apresentar tortuosidade nem apresentar mais de 30 cm de altura.

Souza *et. al.* (2022) realizaram uma avaliação de qualidade de mudas, ao as classificarem em “conforme” e “não conforme”, a fim de verificar a performance de uma plantadora tripla automatizada. Foram observadas 86 mudas não conforme (28,6% do total) e 73% dessas apresentaram tortuosidade elevada como principal parâmetro de impacto. Os autores relatam que esse parâmetro pode ocasionar entupimento dos tubos de plantio e, conseqüentemente, causar falhas no plantio.

2.2 Regressão logística

Beserra *et. al.* (2022) relatam que a regressão logística é um método que foi desenvolvido com o objetivo de descrever os problemas decorrentes do crescimento demográfico do século XIX. Os autores ainda comentam que ela é uma técnica que tem a característica de formar um modelo preditivo para descrever uma variável qualitativa em função de variáveis independentes e é utilizada em diversas áreas do conhecimento. Souza (2013) utilizou a regressão logística para determinar chance de óbito causado pelo câncer de próstata em homens diagnosticados com neoplastia. Mendes & Vega (2011), por sua vez, utilizaram-se da técnica para realizar uma análise ambiental de alteração no padrão do uso do solo no pantanal brasileiro.

Portanto, a regressão logística é usada para determinar a predição de valores que uma variável resposta categórica (normalmente binária) pode assumir. Tal predição é feita a partir de um conjunto de observações de variáveis independentes que podem ser contínuas e/ou categóricas. Em outras palavras, com essa técnica é possível obter a probabilidade de dado evento ocorrer a partir de uma observação aleatória (GONZÁLEZ, 2018).

Após realizar o ajuste dos modelos, é necessário verificar a significância dos coeficientes estimados. González (2018) lista três testes que são os mais utilizados para avaliar o modelo logístico ajustado: Razão da Verossimilhança; Teste de Wald; e Pseudo R^2 de Cox e Snell. No presente trabalho, foi utilizado o Teste de Wald que é o teste padrão disponível no *software* RStudio. Martins (2020) mostra que a estatística do teste é uma razão entre a estimativa de máxima verossimilhança e o erro padrão do estimador para o parâmetro analisado. O autor relata ainda que a estatística apresenta distribuição normal.

Ao utilizar o método de regressão logística, diversos modelos podem ser gerados e é preciso compará-los para poder selecionar o melhor. Emiliano (2009) descreve dois métodos de comparação: o Critério de Informação de Akaike (AIC); e o Critério de Informação Bayesiano (BIC). No presente trabalho foi usado o AIC que, de acordo com o autor, é uma ferramenta de seleção de modelos que pode ser empregada para classificá-los. Os modelos que possuem menor AIC, são considerados melhores.

Ao escolher o melhor modelo ajustado, é necessário interpretar o que os parâmetros β_i representam. Para realizar tal interpretação, é comumente utilizado a razão de chances. De acordo com Souza (2013), a razão de chances pode ser definida como a chance de um evento ocorrer e a chance de o evento não ocorrer. Martins (2020), define a razão de chances como

sendo a comparação entre a probabilidade de sucesso com a probabilidade de fracasso de um dado evento. A interpretação do valor da razão de chances se dá em como a chance de o evento ocorrer ($Y = 1$) é modificada à medida em que são acrescentadas unidades no valor da variável em questão. Além disso, se β_i for positivo, a probabilidade de ocorrência do evento aumenta; se β_i for negativo, essa probabilidade diminui; e, se β_i for igual a zero, a probabilidade não é alterada (MARTINS, 2020).

Por fim, é preciso verificar a qualidade do modelo ajustado. Para isso, no presente trabalho, são utilizadas duas técnicas: a matriz de confusão; e a Curva ROC.

Fernandes *et. al.* (2020) relatam que a matriz de confusão retrata a capacidade preditiva do modelo. Martins (2020) diz que a matriz apresenta os totais de erro e acerto do modelo. Desta Tabela é possível obter os valores de sensibilidade (percentual de ocorrências de interesse, $Y = 1$, que foram corretamente preditas pelo modelo), especificidade (percentual de ocorrências de não interesse, $Y = 0$, que foram corretamente preditas pelo modelo) e acurácia (taxa de acerto) do modelo (FERNANDES *et. al.*, 2020).

A Curva ROC (*Receiver Operating Characteristic*) nada mais é do que uma abordagem gráfica entre os verdadeiros positivos (a sensibilidade, no eixo das ordenadas) e os falsos positivos (1 - especificidade, no eixo das abscissas). Em outras palavras, a Curva ROC é uma representação visual do desempenho do modelo (RODRIGUES, 2020). E, ao analisar a Curva ROC, deve-se observar a área abaixo dela: quanto mais próximo de 1, melhor será o poder discriminativo do modelo (RODRIGUES, 2020). Beserra *et. al* (2022) afirmam que o modelo será mais preciso quando a curva estiver o mais próximo possível do canto esquerdo superior do gráfico.

3 MATERIAIS E MÉTODOS

3.1 A empresa

A Celulose Nipo-Brasileira S.A. – CENIBRA é uma empresa pertencente ao grupo JBP (*Japan Brazil Paper and Pulp Resources Development*) que produz celulose branqueada a partir da fibra curta do eucalipto e possui capacidade de produção de cerca de 1,2 milhão de toneladas de celulose por ano. A fábrica da empresa fica localizada às margens do Rio Doce no município de Belo Oriente, estado de Minas Gerais.

A CENIBRA ainda conta com um viveiro florestal, construído em uma área de 40 mil metros quadrados localizado próximo à unidade industrial da empresa, onde, através da técnica de propagação vegetativa miniestaquia, são produzidas aproximadamente 20 milhões de mudas clonais de eucalipto por ano.

A empresa possui pouco mais de 130 mil hectares de floresta plantada distribuídos na porção leste do estado de Minas Gerais. Além disso, são mais de 100 mil hectares de mata nativa preservada nas formas de Áreas de Preservação Permanente e Reserva Legal.

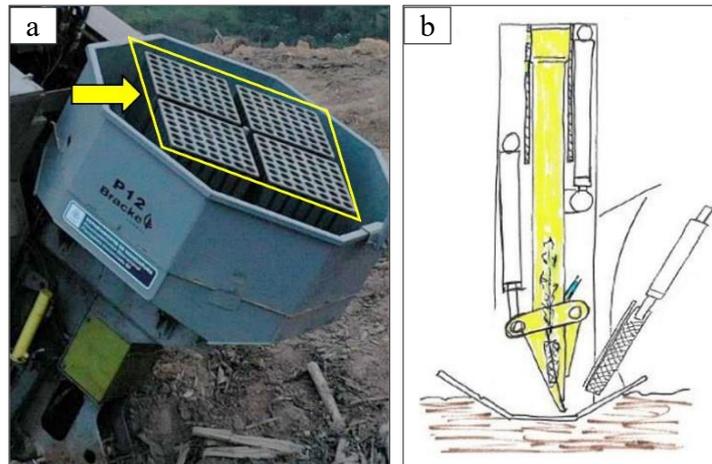
3.2 Mensuração dos parâmetros morfológicos das mudas

No viveiro florestal da empresa, foram avaliadas 196 mudas de dois materiais genéticos de propriedade exclusiva da CENIBRA (392 mudas no total) denominados Clone A e Clone B, clones híbridos de *Eucalyptus grandis* x *Eucalyptus urophylla*, com média de 145 dias de idade.

Em cada muda foram avaliados os seguintes parâmetros: a altura (ALT) em centímetros, o diâmetro do colo (DIAM) em milímetros, a tortuosidade (ANG) do caule em relação ao nível do substrato em graus, o número de pares de folhas (FOL) expandidas no ramo principal, a condição do substrato (SUBS), a região do substrato (POS) onde foi realizada o estaqueamento e a ocorrência de galhos laterais/bifurcação (BIF). Como variável resposta (Y), foi avaliado se a muda ficou retida no magazine¹ (Figura 2a) ou não.

¹ Componente estrutural do cabeçote de plantio onde as mudas que serão plantadas ficam depositadas durante a operação.

Figura 2 – Estruturas do cabeçote: magazine (a); tubo de plantio (b).



Fonte: Do autor (2023).

A altura foi medida com o auxílio de uma régua milimetrada (ELOY *et. al.*, 2013) ilustrada pelas Figura 3a e Figura 3b. Ela foi posicionada na extremidade do tubete ao mesmo nível do substrato, de forma com que fosse possível medir a altura total da muda até o seu ápice. O diâmetro do colo foi determinado com a utilização de um paquímetro (ELOY *et. al.*, 2013) de acordo com a Figura 3c abaixo. Para medir a tortuosidade da muda em relação ao substrato, foi utilizado um transferidor cujo centro foi posicionado sobre o colo da muda (Figura 3d). O valor foi determinado seguindo o prolongamento do caule.

A condição do substrato foi definida como “firme” ou “não firme”. Para isso, foi necessário retirar as mudas dos tubetes e avaliar o estado do substrato naquele momento. As mudas que apresentavam o substrato consistente eram consideradas como “muda com substrato firme”; já aquelas que não possuíam substrato consistente ou com algum defeito tal como esfarelamento, eram consideradas como “muda com substrato não firme” (Figura 3e e Figura 3f).

A região do substrato onde as miniestacas foram inseridas durante o processo de produção das mudas foi definida como “no centro” ou “fora do centro”. Caso o coleto da muda estivesse posicionado de forma deslocada em relação ao centro do tubete (Figura 3h) a muda era considerada como “muda fora do centro”. Caso contrário, ela era considerada como “muda no centro” (Figura 3g).

A ocorrência de galhos laterais ou bifurcação, por sua vez, foi determinada da seguinte forma: caso a muda possuísse algum ramo lateral que apresentasse pelo menos uma folha desenvolvida, ou caso a muda apresentasse bifurcação onde houvesse perda de dominância, a muda era definida como “bifurcada” (a Figura 3i exemplifica o caso de uma muda considerada

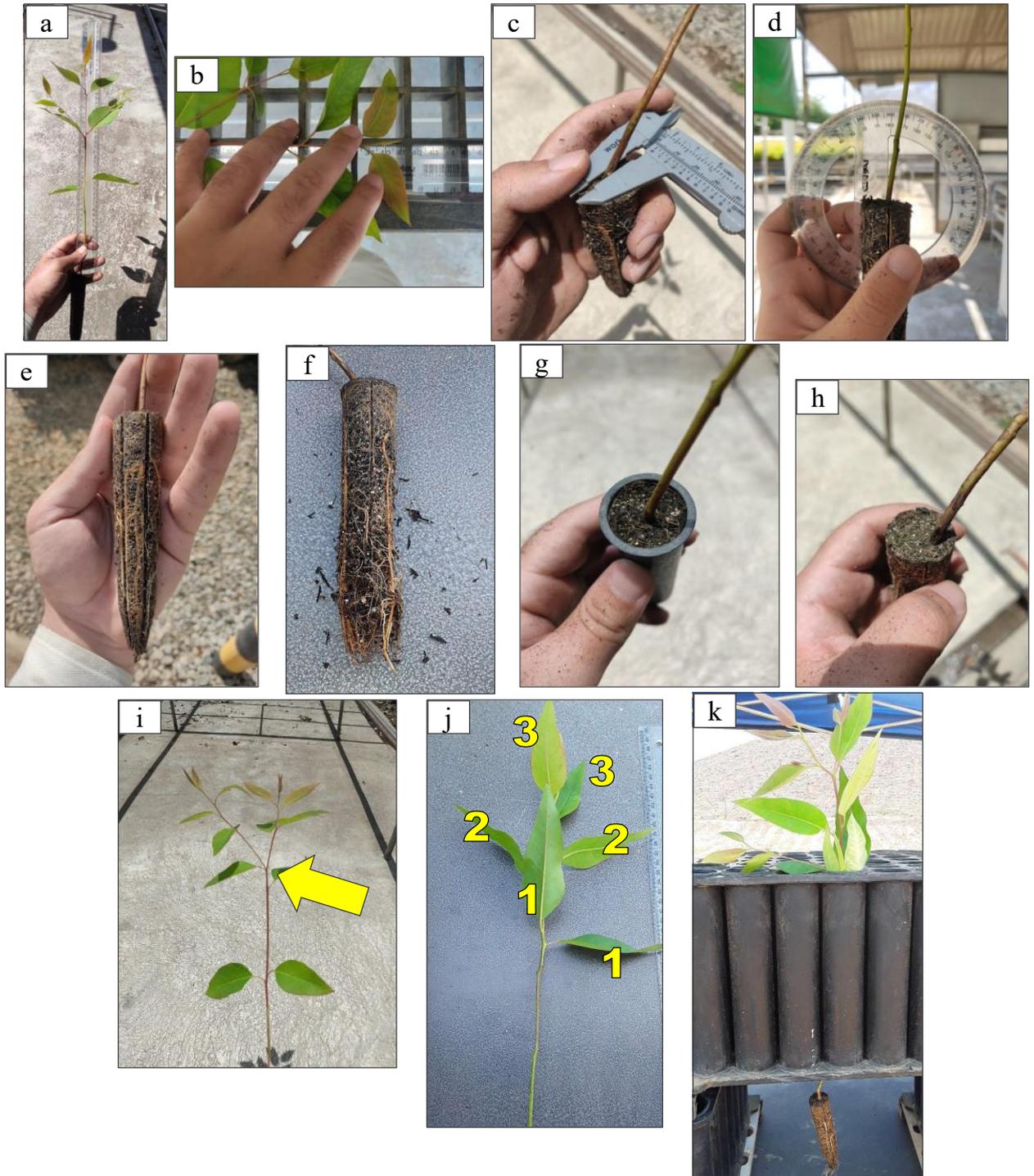
bifurcada. A seta indica a existência de um ramo lateral que apresentou folhas desenvolvidas). Caso contrário, ela era definida como “não bifurcada”.

A determinação do número de pares de folhas foi feita através de contagem. Foram consideradas apenas as folhas que estavam totalmente desenvolvidas do ramo principal, ou seja, folhas de ramos laterais não foram levadas em consideração na contagem (Figura 3j).

Além destes parâmetros, foi mensurada, como variável resposta, o evento de a muda passar ou não passar para o tubo de plantio² (Figura 2b). Para isso, foi utilizado um dos magazines de reposição de mudas com sua extremidade inferior obstruída. A muda foi inserida no magazine e, como durante o processo de plantio ela se desloca por gravidade, a obstrução foi retirada. Para os casos nos quais as mudas deslocaram sem dificuldade, o evento era registrado como “muda passou” e, nos casos que as mudas ficaram retidas no magazine mesmo após a retirada da obstrução (conforme ilustra a Figura 3k abaixo), o evento foi registrado como “muda não passou”.

² Componente estrutural interno do cabeçote de plantio por onde a muda se desloca.

Figura 3 – Parâmetros morfológicos avaliados: altura (a, b); diâmetro do colo (c); inclinação basal (d); agregação do substrato (e, f); posicionamento da muda no tubete (g, h); bifurcação (i); número de par de folhas (j); avaliação do evento de interesse (k).



Fonte: Do autor (2022).

3.3 Registro e análise dos dados

Os parâmetros foram mensurados e registrados no aplicativo ArcGIS Survey123 (ESRI), cuja utilização é bastante difundida na empresa. Para tanto, foi criado um relatório em tal *software* onde foi possível registrar as informações conforme mostra a Figura 4.

Figura 4 – Interface do aplicativo ArcGIS Survey123.

Fonte: Do autor (2022).

Após a coleta dos dados, utilizou-se a versão web do Survey123 (<https://survey123.arcgis.com/>) para criar um arquivo com as informações para cada material genético. O arquivo foi gerado com extensão XLS com o intuito de realizar sua leitura no *software* Microsoft Excel tendo o objetivo de tratar, organizar os dados e facilitar a execução das demais análises estatísticas.

O tratamento da base de dados teve como finalidade realizar uma limpeza na base ao excluir colunas que não teriam utilidade durante a análise estatística tais como informações sobre data da coleta, coluna de identificação de cada registro, coordenadas geográficas, dentre outras. Além disso, a linha de cabeçalho foi renomeada a fim de facilitar o processamento no RStudio.

Como a variável resposta do presente trabalho é do tipo categórica/binária, decidiu-se realizar a análise dos dados seguindo o modelo estatístico de regressão logística binária, uma vez que se visa estimar a probabilidade de o evento Y ocorrer em função dos parâmetros morfológicos citados anteriormente. Segundo Gonzalez (2018) a regressão logística é uma técnica que realiza uma modelagem de previsão utilizando variáveis preditoras a fim de tentar prever o valor de uma variável resposta categórica que geralmente é dicotômica. Mesquita (2014) descreve a regressão logística como uma relação entre as variáveis independentes – que podem ser quantitativas ou qualitativas – e a variável dependente binária. Logo, com um modelo gerado a partir da técnica de regressão logística é possível prever a chance de um dado evento ocorrer a partir de quaisquer observações aleatórias (GONZALEZ, 2018).

4 RESULTADOS E DISCUSSÃO

Foram realizados 392 registros no total. Em 73,2% (287) das ocorrências o evento de interesse (Y=1) ocorreu e em 26,8% (105) delas o evento não ocorreu.

Considerando apenas as observações do Clone A, em 63,3% (124) dos registros o evento (Y=1) aconteceu e em 36,7% (72) deles não houve a ocorrência. Já para o Clone B, o evento ocorreu em 83,1% (163) das ocasiões e em 16,9% (33) ele não foi observado.

Tabela 1 – Resumo dos dados

VARIÁVEL	DESCRIÇÃO	TIPO	UNIDADE
ALT	Altura da muda	Contínua	Em centímetros
DIAM	Diâmetro do coleto	Contínua	Em milímetros
ANG	Inclinação basal	Contínua	Em graus
FOL	Número de par de folhas desenvolvidas	Categórica	1 a 9
BIF	Bifurcação/ramificação	Categórica	0 = Não 1 = Sim
SUBS	Condição do substrato	Categórica	0 = Firme 1 = Não firme
POS	Posição de estaqueamento da muda	Categórica	0 = No centro 1 = Fora do centro
Y	Comportamento da muda no magazine	Categórica	0 = Não desceu 1 = Desceu

Fonte: Do autor (2023).

4.1 Análise generalizada dos dados

4.1.1 Ajuste do modelo

Inicialmente, foi ajustado um modelo geral para os 392 registros considerando que o evento Y ocorra em função de todos os parâmetros mensurados. Os coeficientes deste modelo

1 são indicados na Tabela 2 e os valores p obtidos são provenientes do Teste de Wald. O primeiro modelo apresentou um AIC de 335,11.

Tabela 2 – Estatísticas de ajuste do primeiro modelo generalizado.

	Estimate	Std Error	Z value	Pr(> z)
(Intercept)	11,38255	2,09438	5,435	5,49e-08 ***
ALT	-0,02155	0,03262	-0,661	0,5088
DIAM	-1,08503	0,27174	-3,993	6,56e-05 ***
FOL3	0,24609	0,89192	0,276	0,7826
FOL4	0,03611	0,88965	0,041	0,9676
FOL5	-0,37656	0,90906	-0,414	0,6787
FOL6	0,04526	0,97271	0,047	0,9629
FOL7	1,63875	1,83192	0,895	0,3710
FOL9	-16,33704	882,74389	-0,019	0,9852
BIF [Sim]	-0,22698	0,33216	-0,683	0,4944
SUBS [Nao Firme]	0,09702	0,50703	0,191	0,8484
POS [Fora do centro]	-0,80465	0,31070	-2,590	0,0096 **
ANG	-0,13958	0,01726	-8,087	6,12e-16 ***

Fonte: Do autor (2023).

A partir do modelo 1, é possível observar que os parâmetros DIAM, POS e ANG foram estatisticamente significativos. Portanto, na sequência, foi gerado um novo modelo que levou em consideração apenas esses parâmetros que foram significativos ao nível de confiança de 95%. Os coeficientes obtidos a partir do novo ajuste são mostrados na Tabela 3 e, assim como no modelo 1, os valores p são provenientes do Teste de Wald. O segundo modelo apresentou um AIC de 325.

Tabela 3 – Estatísticas de ajuste do novo modelo generalizado.

	Estimate	Std Error	Z value	Pr(> z)
(Intercept)	10,71679	1,26508	8,471	< 2e-16 ***
DIAM	-1,21894	0,22757	-5,356	8,5e-08 ***
POS [Fora do centro]	-0,70008	0,30054	-2,329	0,0198 *
ANG	-0,13651	0,01651	-8,270	< 2e-16 ***

Fonte: Do autor (2023).

Como o modelo 2 apresentou o menor AIC e um menor número de variáveis, pelo princípio da parcimônia ele foi o escolhido para dar prosseguimento na análise. E, após escolhê-lo, foi gerado as razões de chance (Tabela 4) para cada uma de suas variáveis.

Tabela 4 – Razões de chance para o segundo modelo generalizado.

	OR	2,5%	97,5%	p
(Intercept)	4,5107e+04	4,2749e+03	6,1844e+05	< 2,2e-16 ***
DIAM	2,9554e-01	1,8649e-01	4,5660e-01	8,498e-08 ***
POS [Fora do centro]	4,9654e-01	2,7184e-01	8,8640e-01	0,01984 *
ANG	8,7240e-01	8,4317e-01	8,9970e-01	< 2,2e-16 ***

Fonte: Do autor (2023).

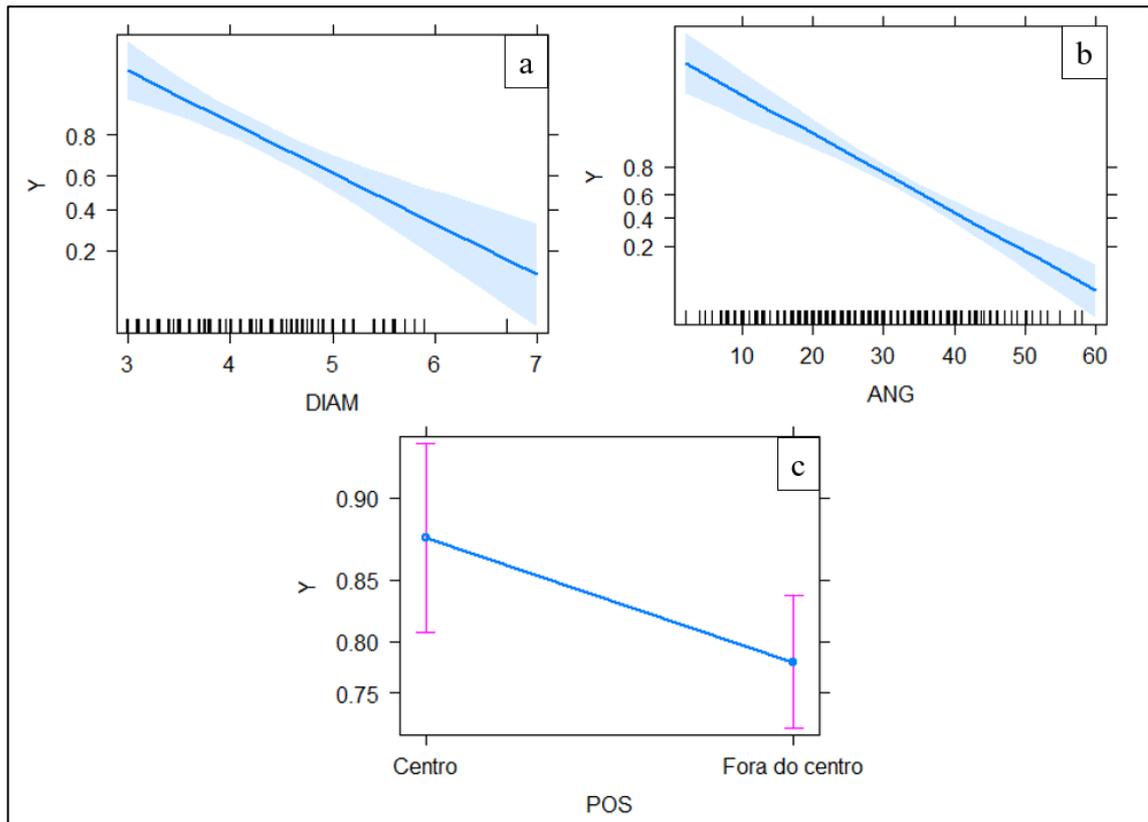
Assim é possível observar que, para o DIAM, à medida que seu valor aumenta com o acréscimo de uma unidade e, considerando que as outras variáveis permaneçam constantes, diminui-se a chance de o evento Y ocorrer em 29,5%.

A mesma interpretação serve para ANG. A cada unidade acrescida em seu valor e, considerando que as demais variáveis sejam fixas, a chance diminui em 87,2%.

Para a variável POS, essa chance é diminuída em 49,6% quando a muda é estaqueada fora do centro do tubete.

Para os três parâmetros, a probabilidade de o evento Y acontecer diminui devido ao fato de os coeficientes (Tabela 3) terem sido negativos. A Figura 5 ilustra tal diminuição. A área ao redor da curva (em Figura 5a e Figura 5b) e as barras de erro (em Figura 5c) representam o intervalo de confiança: quanto maior a área/barra, menor é a confiabilidade no resultado devido a um baixo número de amostras (barras inseridas no eixo das abscissas). A mesma interpretação pode ser utilizada na Figura 7 e Figura 9.

Figura 5 – Efeitos de: DIAM (a); ANG (b); POS (c).



Fonte: Do autor (2023).

4.1.2 Validação do modelo

Para realizar a validação do modelo ajustado, foi levado em consideração a matriz de classificação entre as observações reais e as previsões dele e, em um segundo momento, foi utilizada a curva característica de operação do receptor (Curva ROC).

4.1.2.1 Matriz de classificação

A matriz de classificação foi gerada a fim de verificar a taxa de acerto do modelo ao comparar as previsões obtidas por ele com as observações reais.

A Tabela 5 indica que, considerando o ponto de corte de 0,5, o modelo apresentou uma sensibilidade de 92,3% (265 acertos em relação às 287 ocorrências de $Y=1$) e uma especificidade de 56,1% (59 acertos em relação às 105 observações de $Y=0$). Isso resultou em uma acurácia de 82,6%.

Tabela 5 – Matriz de classificação para o modelo generalizado.

Previsões	Observações reais	
	Não desceu	Desceu
Não desceu	59 *	22 **
Desceu	46 **	265 *

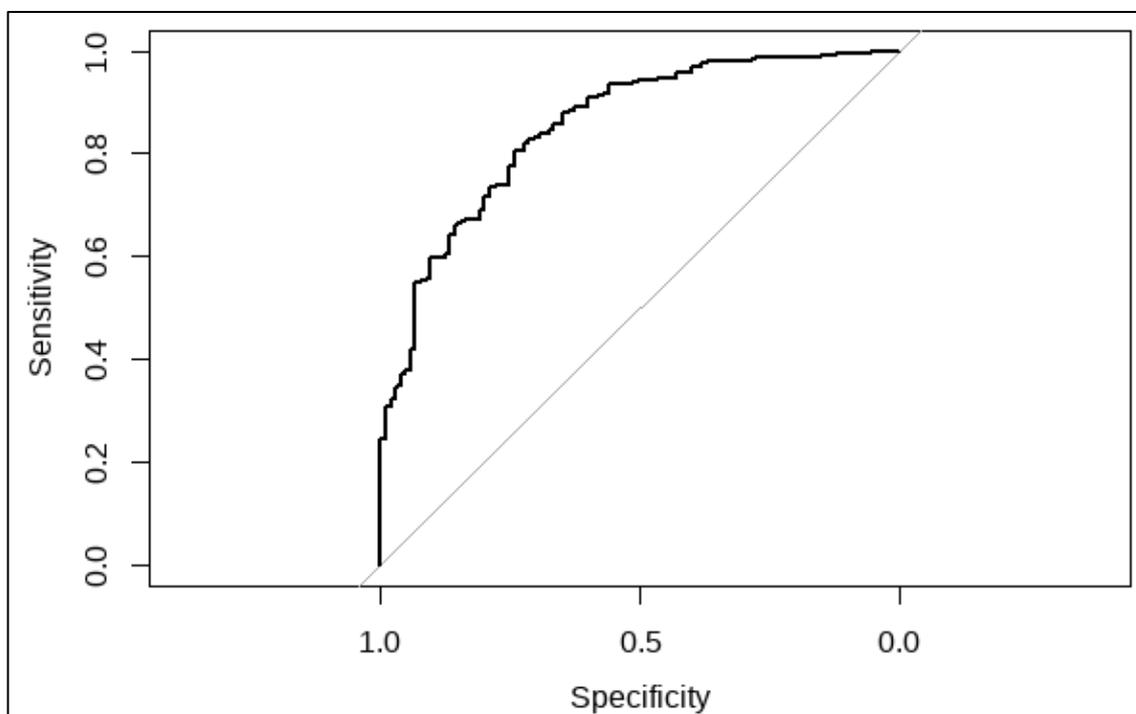
* Previsão correta; ** Previsão incorreta

Fonte: Do autor (2023).

4.1.2.2 Curva ROC

A Figura 6 representa o gráfico da curva característica de operação do receptor (Curva ROC) pois, de acordo com Rodrigues (2020), a Curva ROC é importante para inferir sobre o desempenho do modelo ajustado. Quanto mais próximo de 1 for o valor da área sob a curva, melhor é o desempenho do modelo. Para o modelo ajustado, a área sob a curva foi de 0,8527 o que representa excelente poder de discriminação (MARTINS, 2020).

Figura 6 – Curva ROC para o modelo generalizado.



Fonte: Do autor (2023).

4.2 Análise individualizada dos dados

Ao realizar a análise de forma individual para cada material genético, observou-se uma diferença entre eles. Conforme foi mostrado na seção anterior, de forma geral, os parâmetros que influenciam negativamente no comportamento das mudas durante a atividade de plantio são DIAM, ANG e POS. Porém eles não são comuns aos dois materiais genéticos avaliados.

Para o Clone A, foi notado que após ajustar o primeiro modelo (Y em função de todos os parâmetros mensurados; AIC = 189,78) os parâmetros DIAM e ANG foram estatisticamente significativos ao nível de significância de 95%. Com isso, um novo modelo foi ajustado (Y em função de DIAM e ANG; AIC = 176,58) e, em seguida, foi determinado as suas razões de chance. As estatísticas do novo ajuste são descritas nas Tabelas 6 e 7.

Tabela 6 – Estatísticas de ajuste do modelo do Clone A.

	Estimate	Std Error	Z value	Pr(> z)
(Intercept)	11,78373	1,91623	6,149	7,78e-10 ***
DIAM	-1,43056	0,33540	-4,265	2,00e-05 ***
ANG	-0,15453	0,02345	-6,589	4,43e-11 ***

Fonte: Do autor (2023).

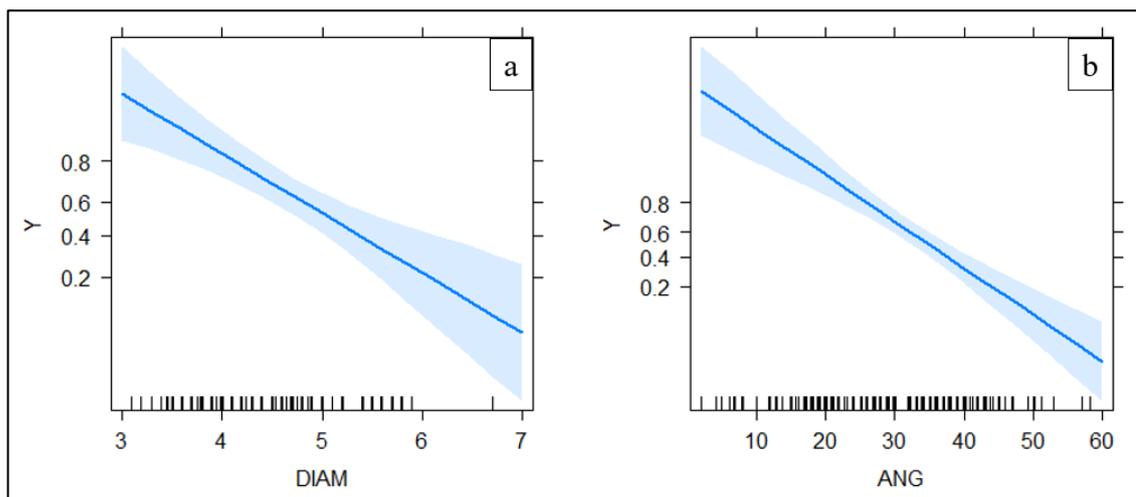
Tabela 7 – Razões de chance para o modelo do Clone A.

	OR	2,5%	97,5%	p
(Intercept)	1,3110e+05	3,8977e+03	7,4443e+06	7,775e-10 ***
DIAM	2,3918e-01	1,1953e-01	4,4940e-01	1,997e-05 ***
ANG	8,5682e-01	8,1536e-01	8,9430e-01	4,433e-11 ***

Fonte: Do autor (2023).

De acordo com a Odds Ratio, para o Clone A, desde que as demais variáveis permaneçam constantes, a chance de o evento ocorrer (Y=1) diminui (os coeficientes da Tabela 6 são negativos; Figura 7) em 23,9% à medida que se acrescenta uma unidade ao DIAM. E, ao acrescentar uma unidade ao ANG (uma vez que as demais variáveis sejam constantes), ela também diminui, porém a uma taxa de 85,6%.

Figura 7 – Efeitos de DIAM (a) e ANG (b) para o Clone A.



Fonte: Do autor (2023).

O modelo ajustado para o Clone A apresentou boa performance visto que a área sob a curva ROC (Figura 8) foi de 0,8667. Além disso o modelo apresentou 85,4% de sensibilidade, 72,2% de especificidade e 80,6% de acurácia. A matriz de classificação é indicada na Tabela 8.

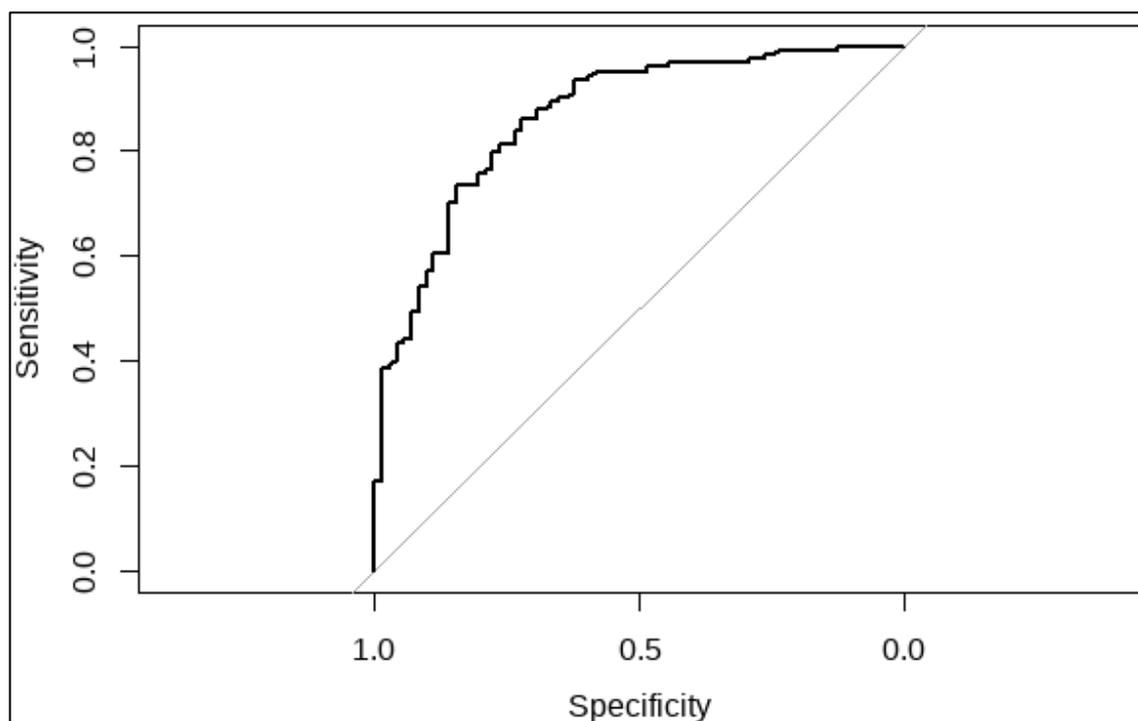
Tabela 8 - Matriz de classificação para o modelo do Clone A.

Previsões	Observações reais	
	Não desceu	Desceu
Não desceu	52 *	18 **
Desceu	20 **	106 *

* Previsão correta; ** Previsão incorreta

Fonte: Do autor (2023).

Figura 8 – Curva ROC para o modelo do Clone A.



Fonte: Do autor (2023).

Para o Clone B, percebeu-se que após realizar o primeiro ajuste de modelo (Y em função de todos os parâmetros mensurados; AIC = 155,35) os parâmetros POS e ANG apresentaram significância ($\alpha = 0,05$). A partir dessa informação, foi realizado o ajuste de um novo modelo (Y em função de POS e ANG; AIC = 148,01) e, em seguida, foi determinado as suas razões de chance. As estatísticas do novo ajuste são descritas nas Tabelas 9 e 10.

Tabela 9 – Estatísticas de ajuste do modelo do Clone B.

	Estimate	Std Error	Z value	Pr(> z)
(Intercept)	5,70222	0,91672	6,220	4,96e-10 ***
POS [Fora do centro]	-1,50296	0,51664	-2,909	0,00362 **
ANG	-0,10744	0,02387	-4,501	6,76e-06 ***

Fonte: Do autor (2023).

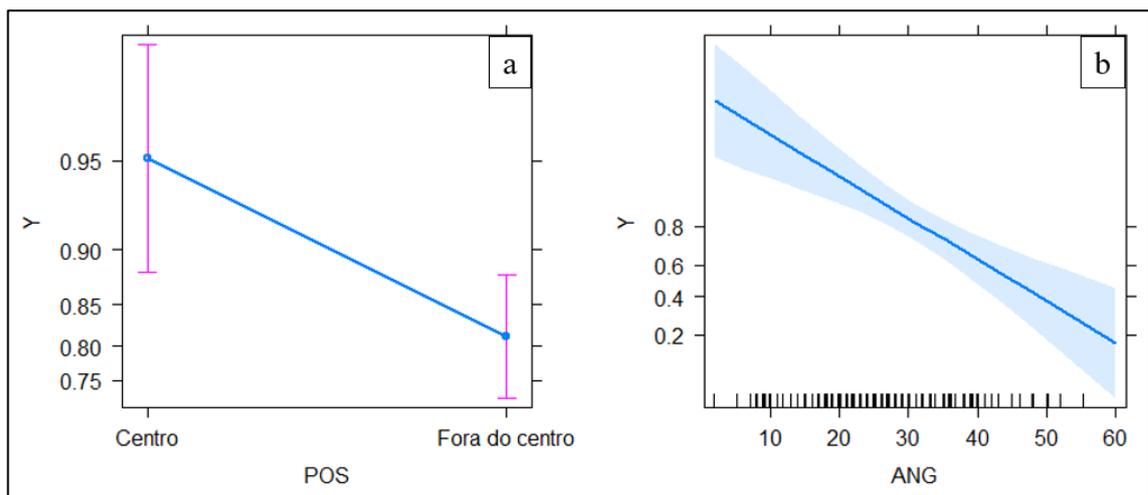
Tabela 10 – Razões de chance para o modelo do Clone B.

	OR	2,5%	97,5%	p
(Intercept)	299,530877	57,818562	2157,1880	4,964e-10 ***
POS [Fora do centro]	0,222472	0,073812	0,5752	0,003625 **
ANG	0,898128	0,854391	0,9388	6,755e-06 ***

Fonte: Do autor (2023).

De acordo com a *Odds Ratio*, para o Clone B, desde que as demais variáveis permaneçam constantes, a chance de o evento ocorrer ($Y=1$) diminui (os coeficientes estimados são negativos; Tabela 9; Figura 9) em 89,8% à medida que se acrescenta uma unidade ao ANG. E, quando POS (a estaquia) é “Fora do centro” e, assumindo mais uma vez que as demais variáveis sejam constantes, ela também diminui, porém a uma taxa de 85,6%.

Figura 9 – Efeitos de POS (a) e ANG (b) para o Clone B.



Fonte: Do autor (2023).

O novo modelo ajustado para o Clone B apresentou boa performance uma vez que sua sensibilidade foi de 98,1%, especificidade de 24,2% e acurácia de 85,7%. Por fim, a área sob a curva ROC (Figura 10) foi de 0,8035. A matriz de classificação é indicada na Tabela 11.

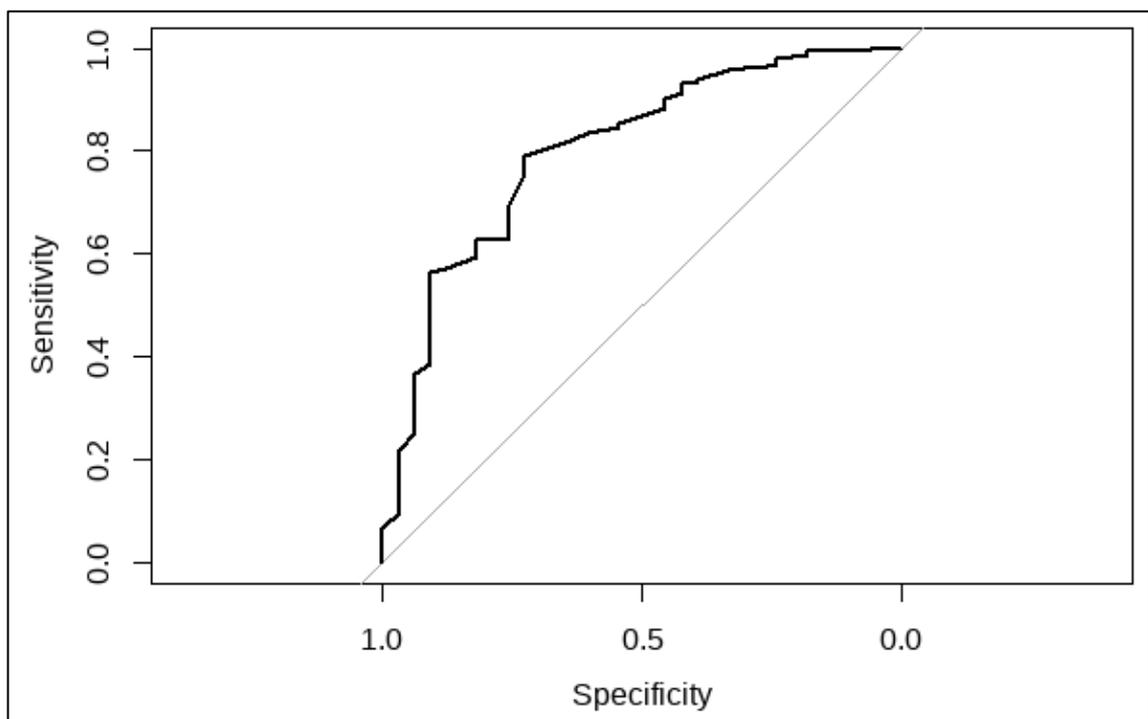
Tabela 11 – Matriz de classificação para o Clone B.

Previsões	Observações reais	
	Não desceu	Desceu
Não desceu	8 *	3 **
Desceu	25 **	160 *

* Previsão correta; ** Previsão incorreta

Fonte: Do autor (2023).

Figura 10 – Curva ROC para o modelo do Clone B.



Fonte: Do autor (2023).

5 CONSIDERAÇÕES FINAIS

De forma geral, a análise dos dados mostrou que o diâmetro do colo, a tortuosidade da muda em relação ao nível do substrato e a posição em que ela é estaqueada no tubete são os parâmetros que influenciam negativamente no fato de a muda não passar pelo magazine do cabeçote sentido ao tubo de plantio. Os demais parâmetros avaliados não apresentaram influência que fosse estatisticamente significativa.

Quando a análise foi feita de forma individual para cada um dos dois clones, foi observada uma diferença entre eles. Porém, a tortuosidade da muda foi um parâmetro morfológico comum entre os dois materiais.

Para o primeiro material genético avaliado, Clone A, foi observado que o diâmetro do colo e a inclinação basal da muda em relação ao nível do substrato foram os parâmetros mais limitantes. Já para o segundo material genético avaliado, Clone B, além da tortuosidade da muda, a posição em que ela foi inserida no tubete durante o processo de produção de mudas influenciou negativamente na ocorrência do evento.

De maneira resumida, à medida que o diâmetro do colo e a inclinação basal aumentam, a probabilidade de a muda descer para o tubo de plantio diminui. O mesmo ocorre quando a muda é estaqueada em uma posição deslocada em relação ao centro do tubete.

Os modelos ajustados apresentaram bom desempenho. O modelo generalizado apresentou acurácia de 82,6%, especificidade e sensibilidade de 92,3% e 56,1%, respectivamente. O modelo gerado para o Clone A apresentou acurácia de 80,6%, especificidade de 85,4% e sensibilidade de 72,2%. Já o modelo gerado para o Clone B indicou acurácia de 85,7%, especificidade e sensibilidade de 98,1% e 24,2%. A área sob a Curva ROC do modelo generalizado, do modelo do Clone A e do modelo do Clone B foi de 0,8527, 0,8667 e 0,8035, respectivamente.

6 CONCLUSÃO

Este trabalho teve como objetivo avaliar alguns parâmetros morfológicos de mudas que são destinadas para o módulo de plantio mecanizado da CENIBRA a fim de identificar quais desses parâmetros poderiam estar interferindo na qualidade da operação uma vez que foram identificadas diversas falhas no plantio. Então, no viveiro da empresa foram avaliadas algumas mudas de dois materiais genéticos que, segundo a recomendação técnica vigente, serão utilizados em uma considerável área de plantio no ano de 2023.

Foi possível observar que a tortuosidade da muda foi um parâmetro comum para os dois materiais genéticos avaliados. Quanto maior é a inclinação basal das mudas, maior é a chance de novas falhas de plantio acontecerem, portanto é preciso destinar mudas que apresentem menor tortuosidade para o módulo de plantio mecanizado. De acordo com os gráficos fatoriais, para o Clone A, mudas com angulação menor que 15 graus apresentam maior probabilidade de passar pelo magazine sentido ao tubo de plantio (Figura 7b). Já para o Clone B, essa probabilidade é atingida em mudas que apresentam inclinação basal inferior a 20 graus (Figura 9b). Logo, durante o controle de qualidade que é realizado no setor de expedição de mudas do viveiro, é necessário incluir uma nova atividade de verificação da tortuosidade da muda para que não haja o risco de que mudas com grande inclinação basal sejam destinadas ao módulo de plantio mecanizado.

A posição onde a muda é estaqueada durante o processo de produção também foi um fator significativo e que influenciou negativamente no desempenho da atividade de plantio mecanizado. Este parâmetro, diferentemente da tortuosidade, não foi comum aos dois clones e interferiu apenas no desempenho do Clone B. As mudas que são estaqueadas no centro do tubete possuem uma chance maior de passarem pelo magazine sentido ao tubo de plantio do que aquelas que são estaqueadas de maneira deslocada em relação ao centro (Figura 9a). Sendo assim, sugere-se que seja passada uma orientação sobre o assunto para a equipe que realiza o estaqueamento no viveiro da empresa.

É sabido que quanto maior for o diâmetro do colo da muda, melhor é a qualidade dela. Então, é preciso avaliar, em trabalhos futuros, o motivo pelo qual a probabilidade de sucesso da atividade reduz quando se utiliza mudas com maiores diâmetros. Acredita-se que tal diminuição possa ser fruto da associação do diâmetro com algum outro parâmetro como a inclinação basal, por exemplo, ou ainda, a idade da muda. Portanto, como sugestão para trabalhos futuros, é necessário avaliar tais relações.

Além disso, é necessário realizar um estudo para poder identificar e entender quais fatores podem estar influenciando e favorecendo o desenvolvimento de mudas com elevada inclinação basal durante o processo de produção no viveiro. E, por fim, é preciso estender o trabalho a mudas de diferentes idades e, além disso, estendê-lo aos demais materiais genéticos utilizados pela CENIBRA para avaliar o comportamento deles também.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANTONANGELO, A.; BACHA, C. J. C. As Fases da Silvicultura no Brasil. **Revista Brasileira de Economia**, Rio de Janeiro, v. 52, n. 1, p. 207–238, 1998.
- BESERRA, R. S. *et al.* Modelagem com regressão logística para análise de concessão de crédito. **Research, Society and Development**, v. 11, n. 7, p. e15211729761, 2022.
- ELOY, E. *et al.* Avaliação da Qualidade de Mudanças de *Eucalyptus grandis* Utilizando Parâmetros Morfológicos. **Floresta**, Curitiba, v. 43, n. 3, p. 373–384, 2013.
- EMILIANO, P. C. **Fundamentos e Aplicações dos Critérios de Informação: Akaike e Bayesiano**. Dissertação (Mestrado) – Universidade Federal de Lavras. Lavras, p. 105. 2009.
- FERNANDES, A. A. T. *et al.* Leia este artigo se você quiser aprender regressão logística. **Revista de Sociologia e Política**, v. 28, n. 74, 2020.
- GONZALEZ, L. de A. **Regressão Logística e suas Aplicações**. Monografia (Graduação) – Universidade Federal do Maranhão. São Luís, p. 45. 2018.
- INDÚSTRIA BRASILEIRA DE ÁRVORES. **Relatório anual 2022: ano base 2021**. São Paulo, 2022.
- INSTITUTO DE PESQUISAS E ESTUDOS FLORESTAIS. **Levantamento do Nível de Mecanização da Silvicultura**. Piracicaba, 2021.
- LAINÉ, T. *et al.* A survey of the Finnish mechanized tree-planting industry in 2013 and its success factors. **Silva Fennica**, v. 50, n. 2, 2016.
- MANNER, J.; ERSSON, B. T. Mechanized tree planting in Nordic forestry: Simulating a machine concept for continuously advancing site preparation and planting. **Journal of Forest Science**, v. 67, n. 5, p. 242–246, 2021.
- MARTINS, A. A. **Aplicação de Análise de Risco de Crédito com o uso das Técnicas de Regressão Logística e Árvores de Decisão**. Monografia (Especialização) – Universidade Federal de Minas Gerais. Belo Horizonte, p. 49. 2020.
- MENDES, C. A. B.; VEGA, F. A. C. Técnicas de Regressão Logística Aplicada à Análise Ambiental. **Revista Geografia**, Londrina, v. 20, n. 1, p. 5-30, 2011.
- MESQUITA, P. S. B. **Um Modelo de Regressão Logística para Avaliação dos Programas de Pós-Graduação no Brasil**. Dissertação (Mestrado) – Universidade Estadual do Norte Fluminense. Campos dos Goytacazes, p. 107. 2014.
- RODRIGUES, C. G. **Avaliação de Desempenho de Modelos de Regressão Logística Multivariada através de Curvas ROC num Estudo de RN de Muito Baixo Peso**. Dissertação (Mestrado) – Universidade do Minho. Braga, p. 131. 2020.

SEREGHETTI, G. C. **Qualidade do Plantio Manual e Mecanizado para Eucalipto e Pinus**. Tese (Doutorado) – Universidade Estadual Paulista “Júlio de Mesquita Filho”. Botucatu, p. 85. 2016.

SOLER, R. R. *et. al.* Avaliação Econômica da Operação de Plantio Mecanizado de Eucalipto em Dois Diferentes Espaçamentos. **Energia na Agricultura**, v. 34, n. 4, p. 462–470, 2019.

SOUZA, A. R. A. De, *et. al.* Análise Estatística do Câncer de Próstata por meio da Regressão Logística. **Rev. Bras. Biom.**, São Paulo, v. 31, n. 3, p. 441–448, 2013.

SOUZA, M. F. *et. al.* Análises Qualitativas de uma Plantadora Tripla Automatizada. **9º Congresso Florestal Brasileiro**, [S. l.], v. 1, n. 1, p. 635-638, 2022.