



MINISTÉRIO DA EDUCAÇÃO
Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Aplicadas
Especialização em Ciência de Dados



Priorização de Atividades de Adubação em Florestas de Eucalipto: Uma Abordagem Orientada a Dados

Patrick Santos Silva

João Monlevade, MG
2024

Patrick Santos Silva

Priorização de Atividades de Adubação em Florestas de Eucalipto: Uma Abordagem Orientada a Dados

Trabalho de conclusão de curso apresentado ao curso de Ciência de Dados do Instituto de Ciências Exatas e Aplicadas da Universidade Federal de Ouro Preto, como parte dos requisitos necessários para a obtenção do título de Especialista em Ciência de Dados.

Orientador: Prof. Dr. Carlos Henrique Gomes Ferreira

João Monlevade, MG

2024

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

S586p Silva, Patrick Santos.
Priorização de atividades de adubação em florestas de eucalipto
[manuscrito]: uma abordagem orientada a dados. / Patrick Santos Silva. -
2024.
39 f.

Orientador: Prof. Dr. Carlos Henrique Gomes Ferreira.
Produção Científica (Especialização). Universidade Federal de Ouro
Preto. Departamento de Engenharia de Produção.

1. Análise de dados. 2. Silvicultura. 3. Tomada de decisão - Método da
pontuação ponderada. 4. Ranqueamento. I. Ferreira, Carlos Henrique
Gomes. II. Universidade Federal de Ouro Preto. III. Título.

CDU 004.652:005.53

Bibliotecário(a) Responsável: Sione Galvão Rodrigues - CRB6 / 2526



FOLHA DE APROVAÇÃO

Patrick Santos Silva

Priorização de Atividades de Adubação em Florestas de Eucalipto: Uma Abordagem Orientada a Dados

Trabalho de conclusão de curso apresentado ao curso de Especialização em Ciência de Dados da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Especialista em Ciência de Dados

Aprovada em 12 de junho de 2024

Membros da banca

Prof. Dr. Carlos Henrique Gomes Ferreira - Orientador (Universidade Federal de Ouro Preto)
Prof. Dr. Fabrício Murai (Worcester Polytechnic Institute)
Prof. Dr. Thiago Augusto de Oliveira Silva (Universidade Federal de Ouro Preto)
Me. Leandro Coelho Dalvi (CENIBRA)

Carlos Henrique Gomes Ferreira, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 26/08/2024



Documento assinado eletronicamente por **Carlos Henrique Gomes Ferreira, PROFESSOR DE MAGISTERIO SUPERIOR**, em 26/08/2024, às 14:40, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0766712** e o código CRC **7AFB9505**.

"Dedico este trabalho à minha família, eterna fonte de inspiração e força, que sempre me incentivou a seguir em frente."

Agradecimentos

Agradeço a Deus e à minha família por todo o apoio durante este período desafiador. Expresso minha gratidão à CENIBRA, em particular ao Departamento de Silvicultura, pela oportunidade concedida. À Universidade Federal de Ouro Preto, em especial ao Professor Carlos, por sua orientação fundamental na elaboração deste trabalho. E aos amigos Felipe Guerra e Gabriel, pelos valiosas contribuições.

*“Os erros que usam dados inadequados
são muito menores do que aqueles que não usam
nenhum dado”. Charles Babbage”*

Resumo

Este trabalho busca preencher uma lacuna ao explorar a aplicação prática da análise de dados nas atividades de silvicultura, destacando seu potencial transformador na gestão florestal moderna. No contexto da CENIBRA, uma empresa multinacional de grande porte, a integração de dados provenientes de diversas fontes e a tomada de decisões eficientes representam desafios significativos. Assim, o objetivo deste estudo é desenvolver um modelo de priorização das atividades de adubação florestal, com base na análise de dados. As bases de dados foram coletadas, pré-processadas e organizadas para a realização das análises em duas etapas. Primeiramente, foi conduzida uma análise exploratória dos dados para identificar padrões e variáveis críticas. Em seguida, foi desenvolvido um modelo de pontuação ponderada para a construção de um plano de adubação orientado a dados, que demonstrou grande potencial na geração de informações para a tomada de decisões na operação de silvicultura. A aplicação do modelo de ranqueamento orientado a dados trouxe mais confiabilidade, segurança e controle ao processo, permitindo uma alocação de recursos mais eficiente e o cumprimento dos prazos estabelecidos.

Palavras-chaves: Análise de Dados, Silvicultura, Tomada de decisão - método da pontuação ponderada, Ranqueamento.

Abstract

This work aims to fill a gap by examining the practical application of data analytics in silviculture activities, highlighting its transformative potential in modern forest management. In the context of CENIBRA, a large multinational company, the integration of data from different sources and efficient decision-making is a major challenge. The aim of this study is therefore to develop a model for prioritizing fertilization measures in the forest based on data analysis. The databases were collected, preprocessed, and prepared for analysis in two steps. First, an exploratory data analysis was carried out to identify patterns and critical variables. Subsequently, a weighted scoring model was developed to create a data-driven fertilization plan, which proved to be very suitable to provide information for decision-making in silvicultural measures. The application of the data-driven scoring model brought more reliability, certainty and control to the process and enabled more efficient resource allocation and adherence to set deadlines.

Keywords: Data Analysis, Silviculture, Decision Making - Weighted Scoring Method, Ranking.

Lista de ilustrações

Figura 1 – Formato das planilhas utilizadas.	13
Figura 2 – Desdobramento de atividades que compõem cada grupo de atividade.	14
Figura 3 – Informações sobre as áreas efetivas e residuais dos “ <i>Talhões</i> ” por regime de manejo.	22
Figura 4 – Distribuição acumulada da Área residual.	23
Figura 5 – Demanda de insumos por hectares (a) e por toneladas (b).	24
Figura 6 – Demanda de insumo ao longo do tempo.	25
Figura 7 – Área residual por grupo de atividade.	26
Figura 8 – Dose média por grupo de atividade.	27
Figura 9 – Distribuição acumulada das atividades por classe de restrição de acesso aos “ <i>Talhões</i> ”.	27
Figura 10 – Distribuição acumulada das atividades por classe.	28
Figura 11 – Distribuição acumulada geral das atividades.	29

Lista de tabelas

Tabela 1 – Pesos utilizados para cada cenário.	17
Tabela 2 – Tabela resultante do processamento dos dados.	20
Tabela 3 – Análise descritiva dos dados.	21
Tabela 4 – Atividades mais e menos prioritárias utilizando os critérios do Cenário 1. . .	30
Tabela 5 – Atividades mais e menos prioritárias utilizando os critérios do Cenário 2. . .	32

Sumário

1	INTRODUÇÃO	1
1.1	Objetivos	4
1.2	Organização do Trabalho	4
2	REVISÃO DA LITERATURA	5
2.1	Silvicultura	5
2.1.1	Adubação	6
2.2	Ciência de Dados	7
2.2.1	Análise Exploratória de Dados	8
2.3	Método da Pontuação Ponderada	9
3	METODOLOGIA	12
3.1	Ferramentas Utilizadas	12
3.2	Construção do Modelo de Pontuação Ponderada	12
3.2.1	Coleta e Pré-processamento dos Dados	13
3.2.2	Análise Exploratória de Dados	15
3.2.3	Método de Pontuação Ponderada	15
4	RESULTADOS	18
4.1	Consolidação da Base de Dados	18
4.2	Análise Exploratória dos Dados	21
4.3	Método da Pontuação Ponderada	29
5	CONSIDERAÇÕES FINAIS	34
	REFERÊNCIAS	36

1 Introdução

Silvicultura é a ciência e a prática de manejar florestas e plantações florestais para atender a diversas necessidades humanas, como produção de madeira, conservação ambiental e recreação (EVANS, 2009). A mecanização da silvicultura representa uma importante evolução nas práticas agrícolas, especialmente no contexto da gestão florestal, em que avanços significativos na automação e mecanização das atividades silviculturais vêm sendo testemunhados (RAMANTSWANA; GUERRA; ERSSON, 2020). A crescente utilização de maquinários especializados, como colheitadeiras, processadoras e tratores florestais, bem como um controle mais sistemático dos processos relacionados a essa atividade, o que já apresenta melhorias substanciais (VENANZI *et al.*, 2023).

Toda essa evolução trouxe consigo uma quantidade substancial de dados que, quando devidamente analisados, podem proporcionar melhorias importantes para as empresas. A tecnologia da informação, em particular, as técnicas de modelagem e análise de dados, têm se apresentado como ferramentas essenciais para melhorar os processos operacionais e promover a sustentabilidade ambiental, como, por exemplo, a gestão das informações relacionadas à atividade, o planejamento das mesmas e o uso eficiente de insumos (VETTORAZZI; FERRAZ, 2000; EVANS, 2009). Nesse contexto, a implantação de mecanismos, como a análise e a ciência de dados, em processos torna-se um instrumento estratégico para extrair informações, identificar padrões e contribuir para a melhoria dos processos envolvidos (RODRIGUES, 2001).

No entanto, acompanhando a tendência de usar dados para planejamento e tomada de decisões, muitas empresas ainda enfrentam desafios substanciais quando se trata de gerenciar, compilar e analisar essas informações de maneira eficiente. Softwares de planilha eletrônica continuam amplamente utilizados para o gerenciamento e, em certo nível, para análise de dados em muitas empresas (WATSON, 2014). À medida que as organizações crescem e os dados se acumulam, muitas vezes ocorre uma fragmentação dos dados em várias planilhas diferentes, mesmo quando elas estão integradas a sistemas e serviços em nuvem, como o Office 365 e o Google Docs (GIBBS; SHANKS; LEDERMAN, 2005). A falta de uma visão holística e integrada das informações prejudica a capacidade de análise e tomada de decisões. Com as informações dispersas, é difícil identificar tendências e padrões que poderiam fornecer orientação estratégica. O resultado é uma lacuna entre as informações disponíveis e a capacidade de acessá-las e utilizá-las de maneira eficaz (ARMSTRONG *et al.*, 2020; JOSHI; KAUSHIK, 2021).

A relevância desse problema também se agrava com a escalabilidade das organizações. À medida que uma empresa cresce e divisões hierárquicas são criadas, a quantidade de informações que ela lida tende a aumentar, tornando os processos mais complexos (JOSHI; KAUSHIK, 2021). Sem uma estrutura sólida para gerenciar essas informações desde o início, o crescimento pode se tornar um agravante. Além disso, não podemos ignorar o fator humano nessa equação. A sobrecarga de trabalho, a frustração e o estresse resultantes da compilação manual de informações podem levar a uma queda na moral da equipe e, eventualmente, a perda de talentos (PÉREZ, 2013). Os profissionais desejam utilizar suas habilidades analíticas e criativas em tarefas de valor agregado, em vez de se perderem em tarefas administrativas e monótonas (BROWN; HESKETH; WILLIAMS, 2004).

Um estudo de 2023 realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) mostra que ainda é pequeno o número de empresas que adotam análise de *big data* e técnicas de inteligência artificial no Brasil (Irene Gomes and Umberlândia Cabral, 2023). Especificamente, estima-se que a análise de *big data* seja utilizada em 23,4% das empresas industriais, referindo-se ao uso de técnicas ou ferramentas de software para coletar, processar e analisar grande volume de dados extraídos da própria empresa ou de outras fontes externas. Já a inteligência artificial é ainda menor, sendo 16,9%, com capacidades de coletar e analisar informações, podendo realizar ações com níveis variados de autonomia para atingir objetivos específicos. Por outro lado, a computação em nuvem foi a tecnologia mais utilizada, alcançando mais de 73%, referindo-se aos serviços de tecnologia de informação e comunicação utilizados pela Internet para acessar um conjunto compartilhado de recursos computacionais configuráveis, entre eles, capacidade de armazenamento. Isso evidencia que, apesar de muitos dados já serem coletados e armazenados por muitas empresas, pouco ainda é explorado para melhorias operacionais.

Ainda, é importante ressaltar que a pesquisa supramencionada é abordada ao nível organizacional, desconsiderando a existência de departamentos e repartições específicas que podem enfrentar desafios ainda maiores. Departamentos não especializados em análise de dados podem ter maior dificuldade em gerenciar e explorar informações valiosas que estão armazenadas, resultando em processos menos eficientes e subutilização de dados. Esses departamentos podem não ter acesso às ferramentas ou ao conhecimento necessário para transformar dados brutos em informações acionáveis, o que poderia proporcionar melhorias significativas na eficiência operacional e na tomada de decisões estratégicas. Assim, a falta de integração e especialização nesses setores sublinha a importância de soluções que centralizem e automatizem a análise de dados em toda a organização.

Para abordar essa questão, é importante considerar soluções tecnológicas e processuais. A centralização de dados, a automação de tarefas, a adoção de sistemas de gerenciamento de projetos e a implementação de relatórios são algumas das medidas que podem ser tomadas. A eficiência na gestão de informações não é apenas uma questão operacional, mas também estratégica. As organizações que conseguem coletar, compilar e analisar informações de maneira eficiente têm uma vantagem competitiva clara ao tomar decisões informadas e ágeis (BRAGA, 2000). Portanto, investir na resolução desse problema é investir no futuro da organização e na maximização de seus recursos mais valiosos, como os seus dados e sua equipe. A eficiência na manipulação de dados operacionais pode contribuir de forma positiva, oferecendo subsídios para melhorar o planejamento das práticas silviculturais e promover a sustentabilidade da operação. Dessa forma, economiza-se tempo e reduz a probabilidade de erros humanos, garantindo maior precisão e confiabilidade nas informações apresentadas (SHABAEV; PITUKHIN, 2021).

Um exemplo dessas soluções é a automatização de relatórios capazes de considerar as necessidades específicas da organização e garantir a compreensão e aceitação por parte dos usuários finais (ARMSTRONG *et al.*, 2020; JOSHI; KAUSHIK, 2021). A integração dos dados e automação de relatórios também fornecem a base para o desenvolvimento de trabalhos futuros, explorando técnicas de análise de dados. A análise de dados por meio de linguagens como Python¹ pode automatizar a extração, transformação e carregamento de dados, bem como a geração de relatórios automaticamente. Além disso, técnicas de análise de dados, como a sumarização, análise exploratória e estatística descritiva, permitem identificar padrões e tendências, auxiliando na definição de prioridades e na tomada de decisões estratégicas (MCKINNEY, 2022). Integrar sistemas de relatórios diretamente a diferentes fontes de dados, ainda que em diferentes departamentos, permite que os relatórios sejam gerados com base em dados sempre atualizados. Por fim, desenvolver modelos capazes de fornecer respostas com os dados mais recentes proporciona consistência e facilita a interpretação.

Neste contexto, este trabalho visa preencher uma lacuna ao explorar a aplicação prática da modelagem e análise de dados nas atividades de silvicultura, com foco em explorar, transformar e analisar informações para auxiliar no suporte à tomada de decisões na gestão florestal moderna. A empresa multinacional Celulose Nipo-Brasileira S/A, mais conhecida como CENIBRA, localizada no estado de Minas Gerais, Brasil, apresenta um cenário complexo devido ao manejo simultâneo de florestas em 461 propriedades próprias. Isso impõe a necessidade de mecanismos integrados que auxiliem na captação de dados provenientes de diversas fontes, no pré-processamento, análise desses e na execução de modelos que priorizem as tarefas a serem realizadas entre as diversas propriedades. Técnicas de modelagem e análise de dados são essenciais para lidar com o grande volume de variáveis envolvidas, possibilitando uma visão precisa e integrada das informações operacionais necessárias para a tomada de decisões estratégicas.

¹ <https://www.python.org/>

1.1 Objetivos

O objetivo geral deste projeto é desenvolver mecanismos para o planejamento e priorização de processos relacionados à silvicultura no âmbito da CENIBRA, usando técnicas de modelagem e análise de dados. Em particular, visa-se a eficiência operacional ao superar a dispersão de dados em várias bases de dados de diversos locais de origem, por meio de uma abordagem automatizada, integrada e analítica para a compilação, modelagem, análise e planejamento das atividades de rotina. Dessa forma, o foco reside em tomadas de decisões mais informadas e uma alocação mais eficiente de recursos no referido contexto.

Especificamente, os objetivos deste trabalho são:

- Integrar diferentes fontes de dados que apoiem a tomada de decisão nos processos relacionados à silvicultura, em particular na adubação de cobertura e de calcário, de forma que seja possível a manipulação em um só banco de dados;
- Definir variáveis que impactam na priorização dos processos relacionados à silvicultura na CENIBRA;
- Estabelecer um modelo matemático que considere as variáveis mais relevantes no processo e seja capaz de auxiliar na definição de prioridades das atividades em questão;
- Desenvolver e implementar relatórios automatizados que forneçam uma visão integrada para a tomada de decisão, incluindo técnicas de sumarização e análise exploratória de dados;
- Avaliar o modelo proposto em cenários de interesse da empresa.

1.2 Organização do Trabalho

O restante do documento está organizado da seguinte forma. O Capítulo 2 apresenta os principais conceitos relacionados a este trabalho. Em seguida, o Capítulo 3 detalha a metodologia do trabalho, enquanto o Capítulo 4 apresenta os resultados. Por fim, o Capítulo 5 oferece as conclusões e direções futuras.

2 Revisão da Literatura

Este capítulo discute os principais conceitos relacionados a este trabalho. Inicialmente, a definição de silvicultura é apresentada. Em seguida, os principais conceitos da análise exploratória de dados, essenciais para entender as variáveis do problema, são explorados. Por fim, apresenta-se o método da pontuação ponderada, uma técnica fundamental para a construção de relatórios que auxiliam na tomada de decisão.

2.1 Silvicultura

A silvicultura é a ciência e a arte de cultivar florestas de forma sustentável para diversos propósitos, incluindo a produção de madeira, celulose, papel, energia e conservação ambiental, entre outros. Esta disciplina desempenha um papel crucial na oferta de recursos naturais renováveis e na manutenção dos ecossistemas. A prática da silvicultura vai além do cultivo e colheita de árvores, mas também envolve um conjunto de técnicas e conhecimentos destinados a garantir que as florestas possam fornecer seus benefícios de maneira contínua e sustentável (ASHTON; KELTY, 2018).

Um dos principais objetivos da silvicultura é a maximização da produtividade florestal, visando melhorar a qualidade da madeira, promover a biodiversidade, conservar o solo e a água, além de mitigar os impactos das mudanças climáticas (PAWSON *et al.*, 2013). Através de práticas bem estabelecidas e manejo adequado, a silvicultura busca equilibrar as necessidades humanas com a preservação dos recursos naturais, garantindo que as florestas continuem a desempenhar suas funções ecológicas essenciais (PERRY, 1998).

Além disso, a silvicultura aborda como as florestas crescem, se desenvolvem e interagem com o meio ambiente. Isso inclui técnicas para plantar, proteger, colher e regenerar áreas florestais. O manejo florestal eficaz envolve uma compreensão profunda dos ciclos de vida das árvores, das dinâmicas dos ecossistemas florestais e das interações entre diferentes espécies de plantas e animais. A silvicultura busca equilibrar as necessidades humanas de madeira, alimentos e outros produtos florestais com a conservação da biodiversidade e dos ecossistemas naturais (JENKINS; SCHAAP, 2018; SINGH, 2024).

O crescimento das árvores é influenciado por uma série de fatores, sendo a nutrição mineral um dos mais importantes. Os nutrientes essenciais, como nitrogênio, fósforo, potássio, cálcio, magnésio e micronutrientes, são fundamentais para o metabolismo e o desenvolvimento saudável das árvores (BHATLA *et al.*, 2018). Esses nutrientes desempenham papéis vitais em processos fisiológicos cruciais, como a fotossíntese, a respiração e a formação de tecidos vegetais. A gestão adequada da nutrição mineral nas florestas pode melhorar significativamente a produtividade e a qualidade da madeira produzida, além de contribuir para a resiliência das florestas contra pragas e doenças (FOELKEL, 2005).

Portanto, a silvicultura é uma disciplina complexa que combina conhecimentos de ecologia, biologia, química e gestão ambiental para promover o uso sustentável das florestas (GREBNER *et al.*, 2021). Ao integrar práticas de manejo sustentável e tecnologias avançadas, a silvicultura moderna busca atender às demandas atuais e garantir que as florestas continuem a fornecer seus inúmeros benefícios às futuras gerações. Essas práticas fortalecem os pilares do desenvolvimento sustentável nas atividades florestais (FLINT; FLINT, 2013).

É importante destacar que diversos fatores exercem influência no processo florestal, abrangendo recursos disponíveis, infraestrutura viária e condições climáticas. Os recursos, tanto financeiros quanto tecnológicos, desempenham um papel crucial na adoção e manutenção de práticas sustentáveis de manejo (MOREIRA; SIMIONI; OLIVEIRA, 2017). A presença de estradas é igualmente essencial, facilitando o acesso às áreas florestais para fins de manejo. Além disso, os elementos climáticos, como temperatura, precipitação e fenômenos extremos, têm impacto direto sobre o desenvolvimento, a vitalidade e a regeneração das florestas, bem como sobre a acessibilidade das vias florestais, tornando necessário considerar tais variáveis no planejamento a longo prazo para a conservação e uso responsável das florestas.

2.1.1 Adubação

Uma atividade chave na silvicultura é a adubação, que consiste na aplicação controlada de nutrientes no solo ou diretamente nas árvores para suprir deficiências nutricionais, estimular o crescimento e melhorar a qualidade da madeira. A escolha dos tipos e das quantidades de fertilizante depende das características de cada solo, das necessidades da espécie florestal e do objetivo de manejo (GONÇALVES *et al.*, 2015).

Existem diferentes tipos de adubação utilizados na silvicultura, incluindo adubação de plantio, realizada no momento do plantio das mudas, adubação de cobertura, durante o crescimento das árvores, e adubação foliar, aplicada diretamente nas folhas para correção rápida de deficiências nutricionais (FERREIRA *et al.*, 2008). A adubação na dose e no momento corretos assegura que esses nutrientes estejam disponíveis quando as árvores mais precisam, especialmente durante períodos críticos de crescimento (SILVA, 2011; GONÇALVES *et al.*, 2015).

O foco da adubação neste estudo está na utilização de fertilizantes minerais, concebidos para fornecer os nutrientes absorvidos pelas plantas no solo e atender às exigências nutricionais elevadas das culturas (REETZ, 2017). A composição dos fertilizantes é determinante para classificar os diversos produtos, pois varia de acordo com a concentração dos três macronutrientes essenciais. Por exemplo, a formulação do fertilizante 13-10-20, amplamente utilizada na CENIBRA, contém 13% de nitrogênio (N), 10% de pentóxido de fósforo (P_2O_5) e 20% de óxido de potássio (K_2O).

O calcário desempenha um papel crucial como corretor de pH do solo. Ao ajustar o pH para valores entre 6,5 e 7,0, cria-se uma condição ideal para que a maioria dos nutrientes fique prontamente disponível para as plantas. Entre as várias formas disponíveis, o calcário dolomítico (com 11% de Mg) é amplamente empregado. Além de sua função principal de ajuste de pH, ele também atua na correção da deficiência de magnésio, oferecendo um benefício adicional para a saúde do solo e o crescimento das plantas (REETZ, 2017; LOPES, 1989).

A aplicação de fertilizantes no momento correto maximiza a sua eficácia e minimiza o desperdício, reduzindo os custos de produção e o risco de lixiviação excessiva ou poluição ambiental causada pelo uso inadequado de fertilizantes. A adubação oportuna pode influenciar diretamente a qualidade da madeira produzida, contribuindo para o desenvolvimento de madeira mais densa, resistente e com características desejáveis para diferentes usos industriais (BRITO *et al.*, 1986). Além disso, uma adubação bem programada faz parte de um manejo florestal sustentável, garantindo a saúde contínua da floresta e sua capacidade de fornecer serviços ecossistêmicos e produtos florestais de forma sustentável ao longo do tempo.

2.2 Ciência de Dados

A **Ciência de Dados** é um campo interdisciplinar que combina técnicas de estatística, computação e conhecimento de domínio para extrair informações valiosas a partir de dados. Ela envolve um espectro completo de atividades, desde a coleta e preparação de dados até a análise avançada e a visualização dos resultados. Dentro desse espectro, a **Análise de Dados** se destaca como uma área focada na aplicação de técnicas estatísticas e computacionais para interpretar e extrair informações úteis dos dados, visando informar decisões estratégicas (MCKINNEY, 2022; BRUCE; BRUCE; GEDECK, 2020).

A Análise Exploratória de Dados (AED) é uma fase inicial e crucial dentro do processo de **Análise de Dados**. Seu objetivo principal é investigar e resumir as características fundamentais dos dados, identificar padrões, detectar anomalias e formular hipóteses preliminares. Essa etapa é essencial para preparar os dados para análises mais complexas e modelagens avançadas, como as realizadas na Ciência de Dados. A AED, portanto, pode ser vista como a fundação sobre a qual se constrói toda a análise subsequente (MCKINNEY, 2022; BRUCE; BRUCE; GEDECK, 2020).

2.2.1 Análise Exploratória de Dados

A análise exploratória de dados (AED) envolve a investigação e a sumarização das principais características de um conjunto de dados. Seu objetivo principal é entender o que os dados podem nos revelar além da modelagem formal ou testes de hipóteses mais avançados (MORETTIN; SINGER, 2022; SAHOO *et al.*, 2019). A AED é essencial na etapa de pré-processamento, pois permite aos analistas identificar padrões, detectar anomalias e testar hipóteses preliminares sobre os dados, facilitando uma compreensão mais profunda antes de avançar para técnicas de análise mais sofisticadas (ROY *et al.*, 2018; DATA *et al.*, 2016; CHAUHAN; SINGH; AGGARWAL, 2021).

A AED pode fornecer uma variedade de resultados úteis e direcionamentos importantes. Entre os resultados comuns estão os resumos estatísticos, que incluem medidas de tendência central (média, mediana e moda), medidas de dispersão (variância, desvio padrão, amplitude e intervalo interquartil) e medidas de associação (covariância e coeficiente de correlação). Essas medidas fornecem uma visão geral das características numéricas dos dados (FERREIRA, 2022).

A distribuição dos dados pode ser representada por várias formas gráficas, como histogramas, box plots, gráficos de densidade, gráficos de barras e gráficos de linhas. Essas representações visuais ajudam a compreender a comparação, distribuição, composição e associação das variáveis nos dados (MORETTIN; SINGER, 2022). Por exemplo, histogramas podem revelar a distribuição dos dados, enquanto box plots destacam a presença de outliers e a dispersão dos dados. Gráficos de densidade ajudam a visualizar a distribuição contínua dos dados, e gráficos de barras e linhas podem ser usados para comparações e tendências ao longo do tempo (PATIL, 2020).

A sumarização de dados, uma tarefa comum na AED, envolve condensar e apresentar dados essenciais de forma concisa (WAGNER, 1998; MASCHI, 2018). Técnicas comuns de sumarização incluem o cálculo de médias, medianas e desvios padrão, bem como a criação de tabelas contendo agrupamentos e ranqueamentos dos dados. A sumarização é essencial para compreender rapidamente a estrutura e as características fundamentais dos dados, permitindo a criação de direcionamentos iniciais que orientam análises mais aprofundadas e tomadas de decisão informadas (WAGNER, 1998).

A normalização dos dados é outra tarefa crítica na AED, especialmente para modelos de aprendizado de máquina e técnicas de ranqueamento. Em muitos modelos, é importante que todas as variáveis estejam na mesma escala para garantir que uma variável não domine as outras devido à sua magnitude. A normalização Min-Max é uma técnica comum utilizada para este fim. Ela ajusta os valores dos dados para um intervalo específico, geralmente entre 0 e 1, garantindo que os dados estejam em uma escala comparável (JIAWEI; MICHELINE, 2006).

A normalização Min-Max é calculada pela seguinte fórmula para cada valor x :

$$x_{norm} = \frac{x - \min(X)}{\max(X) - \min(X)} \quad (2.1)$$

onde:

- x é o valor original que estamos normalizando;
- $\min(X)$ é o valor mínimo de X em todo o conjunto de dados;
- $\max(X)$ é o valor máximo de X em todo o conjunto de dados;
- x_{norm} é o valor normalizado de x , que será escalado para o intervalo entre 0 e 1.

A normalização Min-Max oferece várias vantagens, como a manutenção das relações entre os valores originais e a facilidade de interpretação dos dados normalizados. Ela também melhora o desempenho de algoritmos e modelos que são sensíveis à escala dos dados (HAN; PEI; TONG, 2022).

No contexto deste trabalho, a AED é de extrema importância. A fase inicial de sumarização dos dados permite identificar as variáveis mais relevantes e entender suas interações, facilitando a definição de prioridades para a empresa. Além disso, a normalização dos dados assegura que todas as variáveis consideradas no modelo de ranqueamento estejam na mesma escala, o que melhora a precisão e a eficácia das análises subsequentes. A aplicação de técnicas de AED fornece uma base para a construção de relatórios automatizados que auxiliem na tomada de decisão, contribuindo para uma gestão mais eficiente e informada das operações de silvicultura.

A seguir, apresentamos o método da pontuação ponderada, que é utilizado para priorizar tarefas e facilitar a tomada de decisão no contexto da silvicultura.

2.3 Método da Pontuação Ponderada

O Método da Pontuação Ponderada, também conhecido como *Weighted Scoring Model*, *Weighted Sum Model*, ou *Weighted Criteria Analysis*, tem sido amplamente utilizado como facilitador do planejamento e das tomadas de decisão em diversos setores, como o marketing (SILVA,), a seleção de sistemas de tratamento de efluentes em agroindústrias (LIMA *et al.*, 2018), a seleção de ligas de titânio para aplicação biomédica (PESODE *et al.*, 2023), a avaliação de cenários no setor de geração de energia (RIBEIRO; FERREIRA; ARAÚJO, 2013), e a tomada de decisões em saúde (DIABY; CAMPBELL; GOEREE, 2013), entre outros.

Embora ainda pouco explorado no setor florestal (BLAGOJEVIĆ *et al.*, 2019; LATTE-RINI *et al.*, 2022), esse método possui um enorme potencial para auxiliar na priorização de tarefas e na gestão de operações. A principal vantagem do Método da Pontuação Ponderada na silvicultura é a sua capacidade de lidar com múltiplas variáveis complexas de forma integrada. Este método proporciona uma abordagem estruturada para avaliar diferentes critérios, como custo, tempo, impacto ambiental e urgência, permitindo que as decisões sejam baseadas em uma análise quantitativa robusta. Ao atribuir pesos a cada critério de acordo com sua importância relativa, é possível priorizar tarefas e alocar recursos de maneira mais eficiente. Isso é particularmente benéfico na gestão florestal, onde a eficiência operacional e a sustentabilidade são cruciais para o sucesso a longo prazo. A aplicação desse método pode resultar em uma melhor coordenação das atividades de adubação, colheita e manutenção, otimizando o uso dos recursos disponíveis e minimizando o impacto ambiental.

A gestão de operações envolve inúmeras variáveis que precisam ser priorizadas e direcionadas para alcançar resultados superiores. No entanto, no dia a dia, essas informações não estão prontamente disponíveis para sustentar uma tomada de decisão eficiente, dada a complexidade e a quantidade de fatores impactantes. O método da pontuação ponderada é eficaz para gerar informações que sustentam decisões, especialmente em casos que envolvem múltiplas variáveis. Em resumo, este método calcula a média considerando pesos atribuídos a cada valor, o que é útil quando certos valores têm mais importância ou influência do que outros no cálculo da média final.

Matematicamente, a média ponderada é definida por:

$$\bar{x}_w = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} \quad (2.2)$$

onde:

- x_i são os valores individuais;
- w_i são os pesos atribuídos a cada valor x_i .

Este método é especialmente útil quando se precisa comparar alternativas com diferentes atributos ou características relevantes para a decisão (SILVA, 2012). A utilização da pontuação ponderada permite a criação de um sistema de ranqueamento que pode priorizar tarefas, alocar recursos de maneira mais eficiente e melhorar a gestão das operações. Para utilizar o método de pontuação ponderada, as etapas são as seguintes (SLACK *et al.*, 2009):

1. **Identificação de critérios importantes:** Determinação de quais fatores são relevantes para a decisão ou avaliação em questão. Isso pode incluir variáveis como custo, tempo, impacto ambiental, entre outros.

2. **Atribuição de pesos aos critérios:** Reflete a importância relativa de cada critério na decisão final. Os pesos são atribuídos com base na relevância de cada critério e geralmente somam 100%.
3. **Avaliação das opções com base nesses critérios:** Cada opção é avaliada em relação a cada critério, geralmente através de uma pontuação numérica. Esta avaliação pode ser feita com base em dados históricos, análise de especialistas ou simulações.
4. **Cálculo da pontuação ponderada:** As pontuações para cada critério são multiplicadas pelos respectivos pesos e somadas para chegar a uma pontuação total ponderada. A fórmula matemática é a seguinte:
$$\text{Pontuação Total} = \sum_{i=1}^n (w_i \cdot x_i) \quad (2.3)$$
5. **Tomada da decisão:** Avaliação crítica do resultado e utilização racional do resultado para aplicação no processo operacional. Esta etapa envolve a análise das pontuações ponderadas para tomar decisões informadas e priorizar ações de acordo com as necessidades e objetivos da organização.

No contexto deste trabalho, o método da pontuação ponderada pode fornecer relatórios mais precisos a partir da Análise Exploratória de Dados (AED), onde variáveis são identificadas e mapeadas. Durante a AED, as principais características dos dados são investigadas e resumidas, identificando padrões e outliers que podem influenciar a decisão final (BEHRENS, 1997; MYATT, 2007). A normalização dos dados, como discutido anteriormente, garante que todas as variáveis estejam na mesma escala, permitindo comparações justas e equilibradas.

Aplicar o método da pontuação ponderada no contexto da silvicultura pode ajudar a priorizar tarefas essenciais, como a adubação e a alocação de recursos, com base em critérios como eficácia, custo, impacto ambiental e urgência. Ao integrar dados de diferentes fontes e normalizá-los, a pontuação ponderada pode ser utilizada para criar relatórios que destacam as áreas que precisam de atenção imediata, permitindo uma gestão mais eficaz das operações florestais. A capacidade de gerar relatórios detalhados e precisos a partir da AED e da pontuação ponderada facilita a tomada de decisão, fornecendo informações claras e acionáveis que podem ser utilizadas para otimizar a produtividade e a sustentabilidade das práticas silviculturais.

3 Metodologia

Este capítulo está organizado da seguinte forma: inicialmente, as ferramentas utilizadas no trabalho são descritas e, em seguida, as etapas da construção do modelo são apresentadas.

3.1 Ferramentas Utilizadas

Neste estudo, foram utilizados dados extraídos do software GPF (Gestão de Processo Florestal), uma aplicação central do SAP¹, que integra, organiza e gerencia centralizadamente os dados dos processos florestais. O Google Colab², versão Pro, foi utilizado para a execução de códigos Python³, adequado para aprendizado de máquina e análise de dados, devido à sua facilidade de uso e não requerer instalação local. Já para a construção do trabalho, utilizamos as seguintes bibliotecas:

- **Pandas**⁴: Utilizada para limpeza e tratamento de dados, análise exploratória de dados, consultas e visualização de dados.
- **Sklearn**⁵: Fornece ferramentas simples e eficientes para análise preditiva de dados.
- **Scipy**⁶: Oferece ferramentas para execução de tarefas complexas, baseada na biblioteca NumPy.
- **Seaborn**⁷: Biblioteca de visualização de dados estatísticos, baseada na Matplotlib.

3.2 Construção do Modelo de Pontuação Ponderada

A construção do modelo de pontuação ponderada pode ser dividida em três etapas principais: coleta e pré-processamento dos dados, análise exploratória de dados e construção do modelo.

¹ <<https://www.sap.com>>

² <<https://colab.research.google.com>>

³ <<https://www.python.org>>

⁴ <<https://pandas.pydata.org>>

⁵ <<https://scikit-learn.org>>

⁶ <<https://www.scipy.org>>

⁷ <<https://seaborn.pydata.org>>



Figura 1 – Formato das planilhas utilizadas.

3.2.1 Coleta e Pré-processamento dos Dados

Os dados utilizados neste trabalho foram extraídos de várias fontes dentro da empresa, abrangendo cerca de 1.600.000 observações. A extração dos dados no sistema SAP ocorreu no dia 26/03/2024. As principais fontes de dados incluem:

- **Planilha APONTAMENTO:** Dados das atividades em andamento nos talhões plantados a partir de 2022 até março de 2024, extraídos em formato .xlsx do sistema SAP.
- **Planilha CADASTRO:** Informações detalhadas de talhões, incluindo data de plantio e regime de manejo. Extraídos em formato .xlsx do sistema SAP.
- **Planilha DOSES:** Recomendações de adubação específicas para cada talhão, contendo informações sobre doses e tipos de insumos para cada adubação a ser realizada.
- **Planilha DESCASQUE:** Informações sobre o tipo de colheita realizada em cada talhão (com ou sem descasque).
- **Planilha PRAZO RT:** Prazos para realização das atividades.
- **Planilha RESTRIÇÃO DE ACESSO:** Restrições de acesso aos talhões.

O diagrama na Figura 1 ilustra as diversas planilhas utilizadas para construção da base de dados completa.

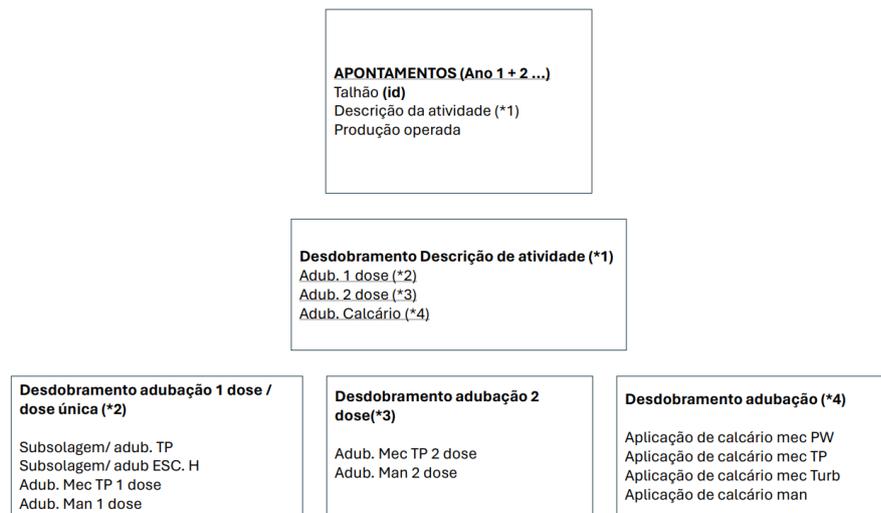


Figura 2 – Desdobramento de atividades que compõem cada grupo de atividade.

Assim, para fins didáticos, utilizamos as informações apenas da região operacional de Santa Bárbara que nos permitirá exercitar os passos a serem replicados pelas demais regiões da Cenibra. Devido à complexidade envolvida e ao grande número de variáveis de diferentes fontes, foi necessário unificar os dados em uma única base de dados. Isso foi feito utilizando a variável “Talhão” como identificador principal, exceto para a planilha PRAZO RT, que utilizou “Data de plantio” como identificador. Durante o processamento, foram criados grupos de atividades (por exemplo, Adubação de primeira dose/dose única, segunda dose, calcário), conforme ilustrado na Figura 2.

Por serem bases de dados obtidas em diferentes fontes e por serem elaboradas por diferentes departamentos, algumas variáveis identificadoras apresentam diferenças nas diversas bases de dados. Os ajustes necessários foram realizados para a padronização dessas variáveis, e que possibilitou o início do processo de unificação das diversas planilhas apresentadas. Esse processo de unificação é crucial para garantir que as diferentes fontes de dados estejam integradas de maneira coerente, permitindo uma análise mais precisa e a construção de modelos que refletem a realidade operacional da empresa. A manipulação dos dados foi realizada com o auxílio das bibliotecas Pandas e Matplotlib, que permitiram editar e ajustar a base de dados conforme os objetivos do estudo. Dessa forma, essa etapa de coleta e pré-processamento dos dados envolveram a integração de dados dispersos na empresa, como as regras de adubação e o regime de manejo dos talhões, facilitando a criação de uma base de dados coesa que suporte a análise exploratória e a construção do modelo de pontuação ponderada.

3.2.2 Análise Exploratória de Dados

O próximo passo foi a Análise Exploratória de Dados (AED). A AED envolveu a investigação e a sumarização das principais características dos dados, utilizando técnicas de resumo estatístico (média, mediana, moda, desvio padrão, etc.) e representações visuais (histogramas, box plots, gráficos de densidade, etc.) para identificar padrões e anomalias. Essa análise, aliada ao conhecimento de domínio, ajudou a selecionar e explorar as variáveis mais relevantes para o modelo. Os resultados detalhados da AED serão apresentados nos diversos capítulos deste trabalho.

3.2.3 Método de Pontuação Ponderada

A Análise Exploratória de Dados (AED), em conjunto com as regras de negócio e informações de domínio, permitiu a seleção das variáveis mais relevantes para a aplicação do método da pontuação ponderada. Este processo de seleção foi crucial para garantir que o modelo considerasse os fatores mais significativos que impactam diretamente as atividades de adubação. Dessa forma, o foco aqui é utilizar essas variáveis selecionadas para construir um modelo robusto que facilite a priorização das atividades, melhorando a eficiência e eficácia das operações florestais.

Para o ranqueamento utilizando o método da pontuação ponderada, consideramos variáveis que impactam as atividades de adubação, como prazo máximo (data limite para execução), acesso ao talhão (dificuldades de acesso em períodos de chuva), dose/ha (dosagem de insumo a ser aplicado), quantidade de talhões (variável criada para quantificação de talhões com pendência de adubação, em um mesmo projeto) e área residual (área restante do talhão sem adubação).

De todas as variáveis consideradas no modelo, o prazo máximo é considerado o mais importante, pois conforme abordado no capítulo anterior, a adubação realizada fora do momento ideal reflete diretamente na produtividade da floresta. As demais variáveis utilizadas também apresentam muita relevância, no entanto, por afetar ou pelo potencial efeito em algum determinado momento, diretamente no cumprimento do prazo máximo. Faremos uma abordagem para cada variável em função do cumprimento do prazo máximo para realização das atividades.

Dentre as variáveis restantes, o " *acesso ao talhão* " apresenta uma forte influência para a realização ou não da atividade de adubação dentro do prazo, principalmente em períodos chuvosos, que nas áreas de atuação da CENIBRA ocorre geralmente de outubro a março. Nesse período há uma limitação de acesso a alguns talhões, e até mesmo a projetos inteiros. Essa limitação está ligada diretamente a topografia e a condição das estradas geradas por longos períodos de chuva. O intuito desse modelo é antecipar a entrada nesses talhões com restrição de acesso antes do início do período das chuvas.

A variável "*Dose/ha*" é bastante relevante quando avaliamos outros fatores que impactam a atividade, como a disponibilidade de mão de obra e de insumo. Quando levamos em consideração o rendimento das atividades, nas atividades com doses maiores são demandadas mais pessoas e, conseqüentemente, mais tempo para a execução, visto que a mão de obra é um recurso limitado. Nesse sentido, em períodos de criticidade em relação a esses recursos apresentados, há a possibilidade de haver uma priorização dos talhões com doses menores.

No entanto, as variáveis "*Quantidade de talhões*" e "*Área residual*" influenciam no cumprimento dos prazos, priorizando talhões com áreas residuais maiores e projetos com maiores números de talhões com adubação pendente. Essa estratégia promove o avanço na produção em área, ela foi traçada com objetivo de aproveitar de forma mais eficiente a mão de obra, visto que o tamanho das equipes são fixas, ou seja, em áreas pequenas ou em áreas maiores, será direcionado o mesmo recurso.

Essa variável "*Quantidade de talhões*" foi criada na base de dados para adicionar um critério que auxiliará no melhor aproveitamento da mão de obra. A quantidade de talhões pendentes possibilita à operação uma melhor alocação de recursos, especialmente em equipes maiores. Dessa forma, uma maior quantidade de talhões amplia a área disponível para a realização das atividades e evita perdas operacionais, como deslocamentos.

A normalização dos dados foi realizada utilizando o método Min-Max, conforme descrito na seção de revisão. Para a variável "*Acesso ao talhão*", foi necessária a codificação dos dados antes da normalização, definindo: Sem restrição: 1, Restrição parcial: 2, Com restrição: 3. A inversão dos dados foi necessária para garantir que as pontuações mais altas reflitam uma melhor performance em todos os critérios, alinhando a direção dos valores com a interpretação desejada. Quando o método de pontuação ponderada considera pontuações mais altas como melhores, mas alguns indicadores (como prazo e dose) são melhores com valores menores, esses indicadores podem ser invertidos. Dessa forma, as pontuações mais altas refletem consistentemente uma melhor performance em todos os critérios. Deste modo é necessário a aplicação dessa técnica para um melhor desempenho do modelo. A equação de ranqueamento é dada por:

$$r = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 \quad (3.1)$$

onde:

- x_1 é o prazo máximo (data limite para execução);
- x_2 é o acesso ao talhão (dificuldades de acesso em períodos de chuva);
- x_3 é a dose/ha (dosagem de insumo a ser aplicado);
- x_4 é a área residual (área restante do talhão sem adubação);
- w_1, w_2, w_3, w_4 são os pesos atribuídos a cada variável x_i , definidos empiricamente.

Diante de todas as variáveis expostas, foram construídos dois cenários para avaliação e utilizado o ranqueamento a partir das variáveis mencionadas. Devido o efeito da variável “Acesso ao talhão” em relação ao prazo, foram criados dois cenários priorizando essa variável. Em cada cenário, os pesos (w_i) foram definidos empiricamente, em consonância com as demandas da empresa, após uma série de testes e ajustes para assegurar que os resultados refletissem os objetivos operacionais na priorização das atividades de adubação. Os resultados detalhados da modelagem, incluindo a eficiência do processamento e a qualidade das informações geradas, são apresentados no próximo capítulo.

O cenário 1, ele não levou em consideração o critério “Acesso ao talhão”, recebendo o peso 0.0 nesse critério. Nesse cenário, o “Prazo máximo” tem o peso igual a 0.6, para que o total dos pesos continuem totalizando 1. Já no cenário 2, tanto o “Acesso ao talhão” quanto “Prazo máximo” receberam pesos 0.3. As variáveis “Área residual”, “Dose”, “Quantidade de talhões” permaneceram com os pesos 0.1, 0.1 e 0.2, respectivamente, nos dois cenários apresentados. Os pesos foram definidos de forma empírica e ajustados com base na análise de diversos resultados obtidos após os ajustes. Esses ajustes foram realizados até se alcançar um resultado realista considerando os cenários e as demandas da empresa, ou seja, um resultado com os ranqueamentos esperados em cada cenário.

A Tabela 1 mostra os pesos atribuídos a cada variável do modelo. Os dois cenários distintos representam estratégias para dois momentos diferentes. O Cenário 1 refere-se a uma abordagem a ser adotada desde o início até a metade do período seco. Por outro lado, o Cenário 2 visa antecipar o início do período chuvoso, permitindo a identificação precoce, principalmente, dos talhões com restrições de acesso.

Tabela 1 – Pesos utilizados para cada cenário.

Variável	Peso no Cenário 1	Peso no Cenário 2
Acesso ao talhão	0.0	0.3
Prazo máximo	0.6	0.3
Área residual	0.1	0.1
Dose	0.1	0.1
Quantidade de talhões	0.2	0.2

4 Resultados

Como apresentado nos capítulos anteriores, o principal desafio deste trabalho é a integração de dados provenientes de múltiplas fontes e a criação de um modelo eficiente para priorizar as atividades de adubação florestal, considerando as várias dimensões e atividades do processo. O objetivo é melhorar a alocação de recursos e garantir o cumprimento dos prazos estabelecidos, além de levar em consideração as variáveis mais impactantes no processo de adubação. Antes de apresentar os resultados do modelo de pontuação ponderada, é fundamental realizar uma análise exploratória dos dados (AED) para identificar os aspectos críticos do processo.

4.1 Consolidação da Base de Dados

Como explicado no Capítulo 3, o pré-processamento dos dados foi realizado de modo a permitir a relação entre as variáveis das diferentes planilhas, possibilitando a unificação dessas bases de dados. Todas as unificações foram feitas utilizando como chave de identificação a variável “*Talhão*”, comum a todas as planilhas utilizadas, exceto a planilha “*PRAZO RT*”, onde a chave de identificação foi a data de plantio, conforme representado na Figura 1.

Após uma análise criteriosa para a seleção das variáveis relevantes, em conformidade com o departamento responsável na CENIBRA, obtivemos uma base de dados única contendo 11 colunas. Foram incluídas na base apenas as variáveis necessárias para a condução das análises, conforme apresentado na Tabela 2. Para o desenvolvimento deste trabalho, as bases de dados e os prazos envolvidos são discutidos e apresentados a partir de uma visão operacional de 07 de junho de 2024.

Considere, por exemplo, a primeira linha da Tabela 2. Inicialmente, precisávamos acessar a planilha “*APONTAMENTO*” e identificar todas as atividades realizadas no “*Talhão*” “*NESBSB00136R24-033*” do grupo de atividades de Calcário (conforme apresentado na Figura 2) e somá-las por meio de tabela dinâmica. Esse procedimento determina a produção operada, valor esse que subtraído da área efetiva do “*Talhão*”, resulta na área residual, ou seja, na parte do “*Talhão*” que continua pendente de realização da atividade prevista. No entanto, essa informação de área efetiva atualizada é obtida dentro da planilha “*CADASTRO*”. Para a adição das demais variáveis, todas as planilhas eram consultadas e inseridas de forma manual para cada atividade prevista, dificultando uma visão simplificada e essencial para a tomada de decisão.

Essa base de dados gerada de forma integrada já disponibiliza de forma imediata todas as informações detalhadas por “*Talhão*”, permitindo o tratamento e a visualização organizada dos dados, o que a torna ainda mais valiosa. Assim, a consolidação de dados realizada na Tabela 2, apresenta os dados essenciais para a elaboração de uma programação de atividades de adubação, incluindo as informações mais complexas de serem determinadas de forma convencional.

Tabela 2 – Tabela resultante do processamento dos dados.

Talhão	Área efetiva (ha)	Atividade	Data de Plantio	Regime	Produção Operada	Área residual	Limite Máximo	Produto Prop.	Dose/ha	Descasque	Prazo max.
NESBSB00136R24-033	12.7	Calcário	2024-02-01	Regeneração	5.9	6.8	2024-08-28	Calcário	1000.0	Sim	82 dias
NESBSB00136P24-007	58.0	1ª dose	2024-02-01	Reforma	22.8	35.2	2024-10-28	10.8.25	500.0	Sim	143 dias
NESBSB00136P24-058	27.7	1ª dose	2024-03-01	Reforma	0.5	27.1	2024-10-28	10.15.20	500.0	Sim	143 dias
NESBSB00128P23-011	7.7	Calcário	2023-12-01	Reforma	2.2	5.5	2024-06-28	Calcário	1500.0	Sim	21 dias
NESBSB00128R23-003	9.0	1ª dose	2023-12-01	Regeneração	4.4	4.6	2024-06-28	13.10.20	600.0	Não	21 dias
NESBSB00159P23-002	34.5	2ª dose	2023-11-01	Reforma	0.0	34.5	2025-03-28	10.8.25	500.0	Sim	294 dias
NESBSB00159P23-001	20.8	2ª dose	2023-11-01	Reforma	0.0	20.8	2025-03-28	13.10.20	500.0	Sim	294 dias
NESBSB00143P23-015	16.3	Calcário	2023-12-01	Reforma	2.8	13.6	2024-06-28	Calcário	1500.0	Sim	21 dias
NESBSB00143P23-018	21.8	Calcário	2023-12-01	Reforma	5.5	16.3	2024-06-28	Calcário	1000.0	Sim	21 dias
NESBSB00143P23-019	10.5	Calcário	2023-12-01	Reforma	2.2	8.3	2024-06-28	Calcário	1000.0	Sim	21 dias
NESBSB00143P23-033	8.4	Calcário	2023-12-01	Reforma	0.2	8.2	2024-06-28	Calcário	1500.0	Não	21 dias
NESBSB00143P23-027	19.0	Calcário	2023-12-01	Reforma	4.4	14.6	2024-06-28	Calcário	2000.0	Sim	21 dias
NESBSB00143P23-036	9.4	Calcário	2023-12-01	Reforma	1.1	8.3	2024-06-28	Calcário	2500.0	Não	21 dias
NESBSB00143P23-035	33.9	Calcário	2023-12-01	Reforma	7.5	26.4	2024-06-28	Calcário	2000.0	Não	21 dias
NESBSB00163P24-078	23.9	Calcário	2024-01-01	Reforma	1.9	22.0	2024-07-28	Calcário	2000.0	Sim	51 dias
NESBSB00163R23-086	50.9	1ª dose	2023-12-01	Regeneração	9.4	41.5	2024-06-28	13.10.20	1200.0	Sim	21 dias
NESBSB00163P24-083	27.9	Calcário	2024-02-01	Reforma	8.1	19.9	2024-08-28	Calcário	1000.0	Sim	82 dias
NESBSB00163P24-083	27.9	1ª dose	2024-02-01	Reforma	24.8	3.1	2024-10-28	10.8.25	500.0	Sim	143 dias
NESBSB00120P24-073	22.0	Calcário	2024-01-01	Reforma	1.0	21.0	2024-07-28	Calcário	2000.0	Sim	51 dias
NESBSB00172R24-016	18.7	Calcário	2024-02-01	Regeneração	12.0	6.6	2024-08-28	Calcário	1000.0	Sim	82 dias
NESBSB00172P24-022	38.1	Calcário	2024-02-01	Reforma	34.5	3.6	2024-08-28	Calcário	1000.0	Sim	82 dias
NESBSB00172P24-022	38.1	1ª dose	2024-02-01	Reforma	25.4	12.7	2024-10-28	13.10.20	400.0	Sim	143 dias
NESBSB00172P24-023	17.9	1ª dose	2024-02-01	Reforma	12.7	5.2	2024-10-28	10.15.20	500.0	Sim	143 dias
NESBSB00172P24-013	17.7	Calcário	2024-02-01	Reforma	14.3	3.4	2024-08-28	Calcário	1500.0	Sim	82 dias
NESBSB00172P24-013	17.7	1ª dose	2024-02-01	Reforma	13.7	4.1	2024-10-28	10.8.25	500.0	Sim	143 dias

Tabela 3 – Análise descritiva dos dados.

	Área efetiva (ha)	Produção Operada	Área residual	Dose/ha	Prazo max.	Quantidade de talhões
Contagem	36.0	36.0	36.0	36.0	36 dias	36.0
Média	26.3	11.8	14.5	1116.7	91 dias	6.2
Min	7.7	0.0	1.7	400.0	21 dias	1.0
25%	17.4	2.1	5.1	500.0	21 dias	5.0
50%	24.2	7.7	8.3	1000.0	82 dias	7.0
75%	34.7	17.9	21.2	1500.0	143 dias	8.0
Max	58.0	51.7	41.5	2500.0	294 dias	9.0
Desvio padrão	14.2	12.6	11.4	562.4	68 dias	2.2

4.2 Análise Exploratória dos Dados

A análise exploratória dos dados foi realizada em duas etapas distintas: uma análise descritiva inicial seguida de uma análise detalhada utilizando Funções de Distribuição Acumulada (CDFs) e intervalos de confiança (ICs). Cada etapa possui um propósito específico e contribui de forma complementar para uma compreensão aprofundada dos dados.

A análise descritiva inicial visa fornecer uma visão geral das características básicas dos dados. Nesta etapa, são calculadas estatísticas descritivas como média, mediana, mínimo, máximo, desvio padrão e quartis para cada variável relevante. Essas estatísticas oferecem uma compreensão rápida e abrangente da centralidade, dispersão e distribuição das variáveis, além de ajudar na identificação de possíveis valores atípicos que possam influenciar negativamente as análises subsequentes. A Tabela 3 apresenta um resumo das estatísticas descritivas dos dados analisados.

Os dados mostram variações significativas em todas as categorias analisadas. A média da área efetiva é 26,3 hectares (ha), com valores variando de 7,7 ha a 58,0 ha e um desvio padrão de 14,2 ha. A produção operada tem uma média de 11,8 e uma variabilidade alta, com um desvio padrão de 12,6. A área residual média é 14,5 ha, também mostrando grande variação. A dose média por hectare é 1116,7, variando de 400,0 a 2500,0 com um desvio padrão de 562,4. O prazo máximo tem uma média de 91 dias, variando entre 21 e 294 dias, e a quantidade média de “Talhões” com pendências de adubação em um mesmo projeto é 6,2, com valores variando de 1,0 a 9,0, com um desvio padrão de 2,2.

Essa análise descritiva destaca a dispersão e a variabilidade dos dados, oferecendo uma visão geral das distribuições e indicando a presença de possíveis outliers ou valores extremos que podem requerer investigação adicional. As medidas centrais como a média e a mediana, com os quartis e o desvio padrão, ajudam a entender melhor os padrões e a dispersão dos dados, proporcionando uma base para futuras análises mais detalhadas ou tomadas de decisão informadas.

A limitação de recursos faz com que seja necessário um planejamento estratégico para o cumprimento pleno das atividades dentro dos prazos pré-estabelecidos pela empresa, principalmente nas atividades de adubação. Esses prazos são definidos a partir de estudos científicos que mostram o ganho em produtividade e qualidade da madeira quando se adotam práticas silviculturais dentro da janela de maior assimilação de nutrientes pelas plantas (RAIJ, 1997). Por esse motivo, uma das grandes preocupações das empresas do setor agrícola e florestal é o cumprimento rigoroso dos prazos. Em termos de recursos, uma das principais informações a ser analisada é a área pendente em relação à área efetiva para a realização das atividades, tratada neste trabalho como área residual. Com essas informações, é possível realizar o planejamento das atividades, tendo como referência a quantidade exata de área em cada atividade. Com esse nível de detalhamento, é possível determinar a quantidade de horas ou dias (HH/dia) necessárias para a realização das áreas pendentes.

Dessa forma, a próxima etapa envolve uma análise mais detalhada das distribuições destes valores utilizando Funções de Distribuição Acumulada (CDFs) e intervalos de confiança (ICs). Essas técnicas permitem uma compreensão mais profunda da distribuição dos dados e das relações entre variáveis, facilitando a identificação de padrões e tendências que não são imediatamente evidentes na análise descritiva. Os gráficos das Figuras 3a e 3b apresentam, respectivamente, a média da área efetiva dos “*Talhões*” e a média da área residual dos “*Talhões*”. Além disso, o intervalo de confiança de 95% é apresentado nas barras.

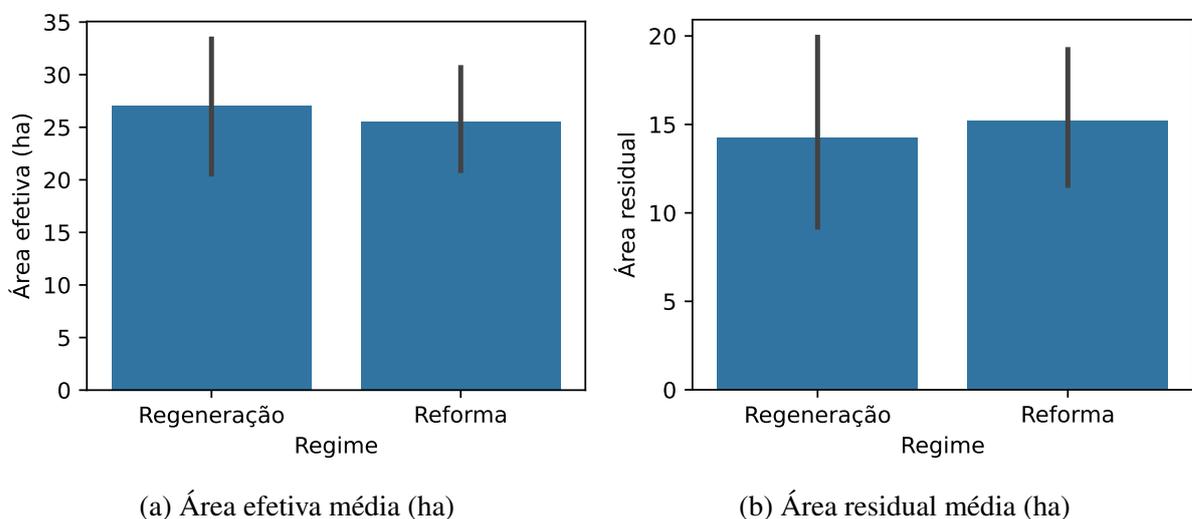


Figura 3 – Informações sobre as áreas efetivas e residuais dos “*Talhões*” por regime de manejo.

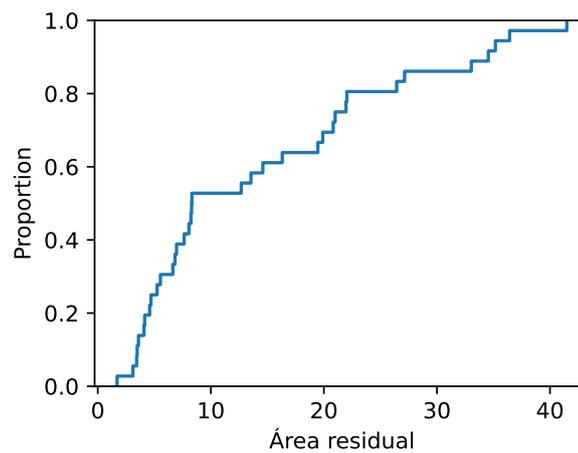


Figura 4 – Distribuição acumulada da Área residual.

Observa-se que a área residual média é substancialmente menor em comparação com a área efetiva. Os intervalos de confiança (ICs) de 95% para ambos os regimes de manejo se sobrepõem significativamente, indicando que não há uma diferença estatisticamente significativa entre as áreas efetivas e residuais quando comparadas entre os regimes de regeneração e reforma. A amplitude do eixo Y na área residual é muito menor do que na área efetiva, refletindo a menor variabilidade das áreas residuais. Isso ocorre devido às limitações de declividade do terreno, que impedem a mecanização em certas áreas. Em “*Talhões*” de regeneração e reforma, a área residual média está em torno de 15 hectares, representando cerca de 55% da área total dos “*Talhões*”. Esse alto percentual de área residual indica um desafio significativo na realização completa das atividades de adubação, especialmente nas áreas onde a mecanização não é viável devido à inclinação do terreno. Atualmente, apenas 75% das áreas com potencial de mecanização estão sendo aproveitadas, o que sugere que uma revisão das práticas de mecanização pode ser necessária.

A Figura 4 mostra a distribuição acumulada das áreas residuais dos “*Talhões*”. Observa-se que a maioria dos “*Talhões*” tem uma área residual relativamente baixa. Cerca de 60% dos “*Talhões*” têm uma área residual inferior a 10 ha. Isso indica que a maioria das atividades pendentes está concentrada em “*Talhões*” com menores áreas, o que pode facilitar a alocação de recursos e o planejamento operacional. À medida que avançamos no eixo X, percebemos que a proporção de “*Talhões*” com áreas residuais maiores (acima de 20 ha) é significativamente menor. Apenas cerca de 20% dos “*Talhões*” têm áreas residuais que variam entre 20 e 40 ha. Esta distribuição sugere que, apesar de haver alguns “*Talhões*” com áreas residuais consideráveis, a maioria das atividades de adubação pendentes está distribuída em áreas menores. Isso pode permitir uma maior flexibilidade no planejamento e execução das atividades, além de uma potencial redução no tempo necessário para a conclusão das tarefas.

A perda de áreas potenciais de mecanização não só agrava a questão da escassez de mão de obra, mas também aumenta a pressão para cumprir os prazos das atividades, transferindo essa demanda para as equipes manuais. Essa análise destaca a importância de intervenções rápidas e embasadas, possibilitando a atuação imediata antes que a operação saia do “*Talhão*”, garantindo assim a eficiência e eficácia no cumprimento dos prazos. A combinação dessas informações permite um planejamento mais robusto, ajudando na alocação do uso de recursos e melhorando a produtividade das operações florestais.

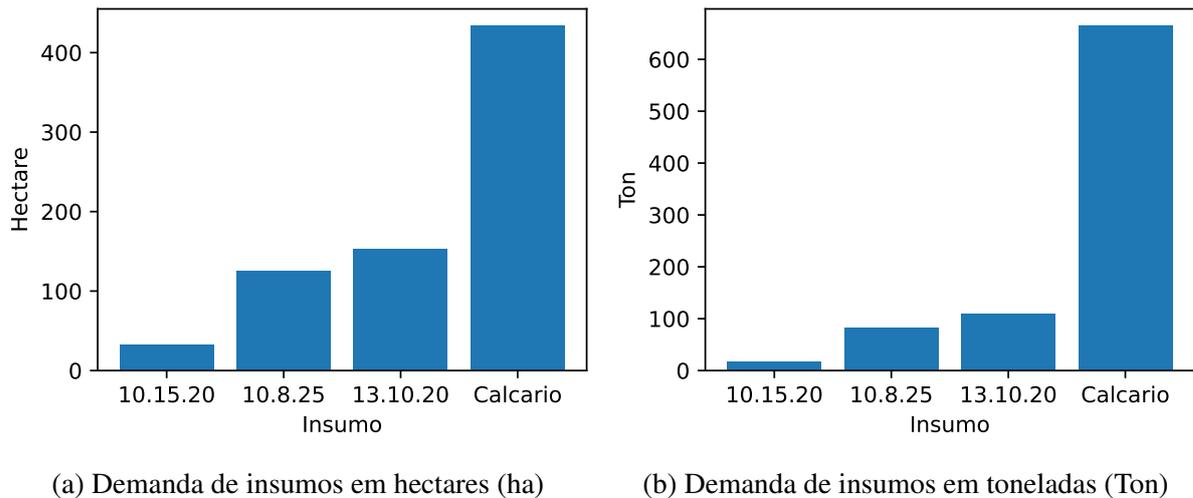


Figura 5 – Demanda de insumos por hectares (a) e por toneladas (b).

Outro fator de suma importância a ser considerado no processo de adubação é a disponibilidade de insumos na quantidade e no momento corretos. A indisponibilidade desses insumos, seja fertilizante ou calcário, reduz a eficiência da operação ao interromper a sequência do planejamento das atividades. Em alguns momentos, isso gera a necessidade de reprogramação de alguns “*Talhões*” e a antecipação de atividades em “*Talhões*” não prioritários. Essas mudanças refletem negativamente no cumprimento dos prazos exigidos e, conseqüentemente, na produtividade das florestas.

Nesse contexto, as Figuras 5a e 5b mostram, respectivamente, as distribuições da demanda de insumos em hectares e em toneladas. Essas informações são fundamentais para a elaboração de um plano de alocação de insumos adequado. Analisando as distribuições dos gráficos na Figura 5, observamos que o calcário apresenta a maior demanda, tanto em hectares quanto em toneladas. A Figura 5a mostra que a demanda de calcário abrange mais de 400 hectares, enquanto a demanda das formulações 10.15.20, 10.8.25 e 13.10.20 é significativamente menor. Essa tendência se repete na Figura 5b, onde a demanda de calcário em toneladas é a mais alta, chegando a mais de 600 toneladas.

Essas observações destacam a importância do calcário no processo de adubação, especialmente por suas funções críticas na correção do solo e na otimização da absorção de nutrientes pelas plantas, essenciais para a produtividade das florestas de eucalipto, e sendo o insumo mais demandado em termos de área e peso. A demanda elevada por calcário pode ser atribuída às suas funções críticas na correção do solo e na otimização da absorção de nutrientes pelas plantas, essencial para a produtividade das florestas de eucalipto. A comparação entre hectares e toneladas reforça a necessidade de um planejamento logístico detalhado para garantir que a quantidade correta de insumos esteja disponível no momento certo. Isso é vital para evitar interrupções na sequência das atividades de adubação e para garantir que os prazos sejam cumpridos de forma eficaz. A análise detalhada das demandas de insumos permite identificar períodos de pico de demanda e preparar a logística para atender essas necessidades, minimizando o risco de atrasos operacionais e maximizando a eficiência das operações.

Além disso, é essencial avaliar a demanda desses insumos ao longo do tempo para planejar o fornecimento e a logística adequadamente. Na Figura 6, conseguimos visualizar a demanda de formulações ao longo de todo o período abordado no estudo.

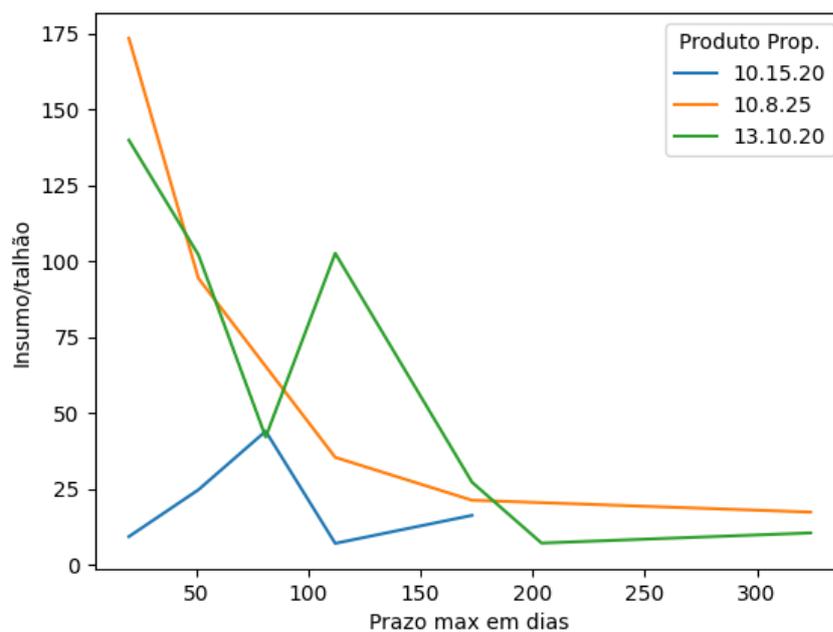


Figura 6 – Demanda de insumo ao longo do tempo.

É possível observar que a demanda por insumos varia consideravelmente ao longo do tempo. Observa-se uma demanda elevada para as formulações próximas ao 25º e ao 75º dia, com um novo pico próximo ao 100º dia. Após esse período, a demanda para esses “Talhões” reduz linearmente, encerrando próximo aos 180 dias. Cada formulação apresenta um comportamento distinto. As formulações 13.10.20 e 10.8.25, por exemplo, têm picos iniciais de demanda, mas a 13.10.20 alcança um pico pouco antes do 50º dia, com aproximadamente 50 toneladas, antes de reduzir gradativamente. Já a formulação 10.15.20 mostra um comportamento diferente, com uma demanda menor e mais concentrada em um período específico, não ultrapassando os 180 dias.

Essas informações sobre o comportamento do consumo de insumos são cruciais para eliminar a indisponibilidade de insumos como um dos fatores que interferem no cumprimento dos prazos das adubações. Com a previsão precisa da demanda, é possível planejar o fornecimento dos insumos necessários, garantindo que estejam disponíveis no momento certo. Esse planejamento robusto permite uma melhor gestão dos recursos, garantindo a eficiência operacional e o cumprimento dos prazos estabelecidos.

Informações sobre o regime também são úteis para direcionar as estratégias operacionais da empresa. Em determinadas situações críticas, ações podem ser necessárias para priorizar alguma das demandas. Nesse momento, a visualização do cenário atualizado de pendências pode reforçar a decisão de priorização de algumas atividades. Distribuições acumuladas de cada atividade pendente são úteis para o direcionamento dos esforços e definição da estratégia operacional a ser adotada.

Na Figura 7, é possível visualizar que a demanda de calcário representa aproximadamente 55,70% dos 520,32 hectares pendentes de realização, superando os 34,69% e os 9,61% de adubação de 1ª dose e 2ª dose, respectivamente. Essa análise destaca a importância de priorizar as atividades de aplicação de calcário, dado seu grande impacto na área total pendente.

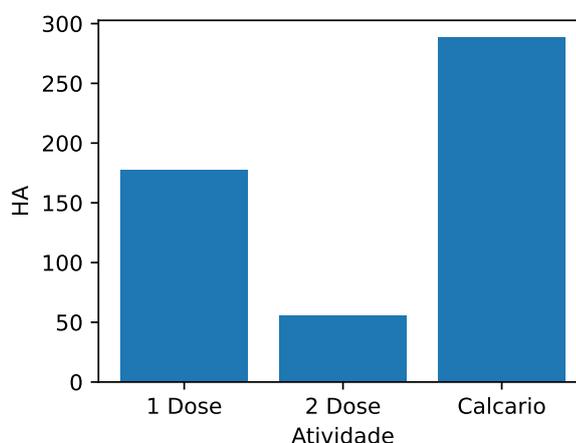


Figura 7 – Área residual por grupo de atividade.

Para a definição das estratégias, além da área residual, informações como a dose por hectare (“*Dose/ha*”) são de suma importância. A dose influencia diretamente no rendimento da atividade, tornando crucial a obtenção dessas informações para uma decisão mais assertiva. Na Figura 8, observamos a dose média por grupo de atividade, com seus respectivos desvios padrões.

A análise dos gráficos revela que a aplicação de calcário requer uma dose significativamente maior em comparação às doses de adubação (1ª e 2ª dose). A dose média de calcário é de aproximadamente $1500 \text{ Kg} \cdot \text{há}^{-1}$, enquanto as doses de adubação são substancialmente menores. A dose de 1ª dose está em torno de $600 \text{ Kg} \cdot \text{há}^{-1}$ e a de 2ª dose em torno de $500 \text{ Kg} \cdot \text{há}^{-1}$. Essas informações são essenciais para a priorização das atividades, pois uma maior dose de insumo pode implicar em maior tempo e recursos necessários para a execução da atividade. Além disso, ao considerar os prazos para a realização das atividades, a análise da dose média, juntamente com a área residual, fornece uma base sólida para a alocação eficiente de recursos e o cumprimento dos prazos estabelecidos. Portanto, a visualização e análise detalhada dos dados de área residual e dose média são fundamentais para a definição de estratégias operacionais eficientes, garantindo que as atividades prioritárias sejam realizadas dentro do prazo, alocando de forma mais eficiente o uso dos recursos disponíveis.

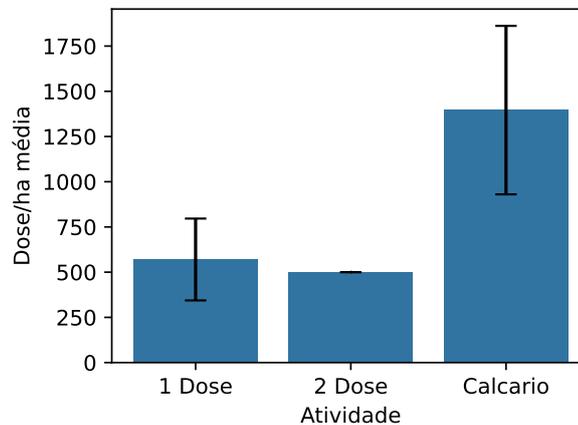


Figura 8 – Dose média por grupo de atividade.

Outro ponto a ser considerado é a condição de acesso aos “*Talhões*”, que compromete fortemente a dinâmica da operação. Por isso, é necessário considerar os “*Talhões*” de difícil acesso e, se necessário, antecipar a entrada, para não comprometer os prazos.

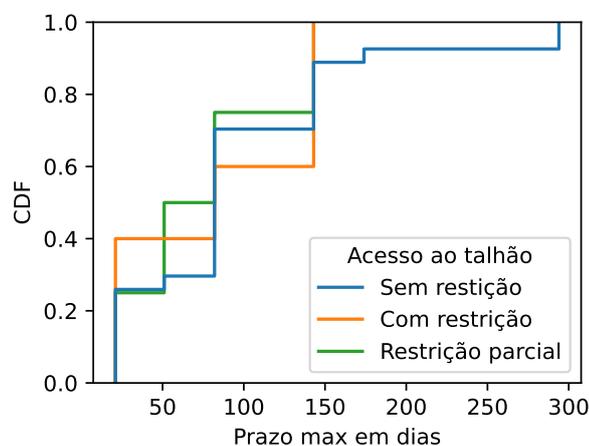


Figura 9 – Distribuição acumulada das atividades por classe de restrição de acesso aos “*Talhões*”.

A Figura 9 apresenta o gráfico de Distribuição Acumulada (CDF) para ilustrar a ocorrência de “*Talhões*” com restrição de acesso ao longo de todo o período. Observa-se que aproximadamente 60% dos “*Talhões*” com restrição de acesso estão nos primeiros 80 dias, outros 25% estão distribuídos entre o 80° e o 150° dia e os 15% restantes a partir do 150°. Essas informações, quando relacionadas com as condições climáticas do período, permitem uma previsão ainda mais detalhada. Com esse embasamento, é possível criar e definir um plano de antecipação de algumas atividades, garantindo que, se realizadas no momento inicialmente planejado, não sejam inviabilizadas pela falta de acesso.

A partir das informações apresentadas no gráfico, é possível observar que apenas 10% das atividades estão planejadas para serem realizadas entre o 180° e o 300° dia. O restante está planejado para ocorrer antes desse período, exigindo uma avaliação da disponibilidade de mão de obra para a realização dessas atividades concentradas nesse intervalo. Com essas informações, será possível identificar a necessidade de reprogramação de algumas atividades em “*Talhões*” sem restrição de acesso e tenham prazos de realização após o 100° dia.

Esse mesmo formato gráfico pode ser aplicado às inúmeras outras variáveis do processo. Como nesse exemplo (Figura 10), que mostra detalhadamente a distribuição acumulada das atividades ao longo do tempo, bem como o momento em que elas vão disputar o mesmo recurso. As atividades de adubação de 1ª dose e calcário estão previstas dentro dos primeiros 200 dias, sendo que 95% do calcário está previsto para conclusão até o 90° dia, não competindo mão de obra com a atividade de 1ª dose. Apenas 20% das atividades de 1ª dose estão programadas para esse período, a porção significativa dessa atividade está com prazo final previsto aos 150 dias, consequentemente com o planejamento correto das operações, dispõe-se de um intervalo de aproximadamente 60 dias para a realização desses 80% das atividades de 1ª dose. Já as adubações de 2ª dose estão concentradas próximo ao 295° dia.

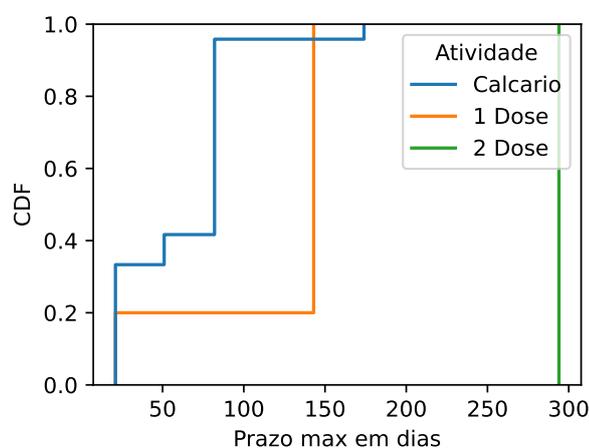


Figura 10 – Distribuição acumulada das atividades por classe.

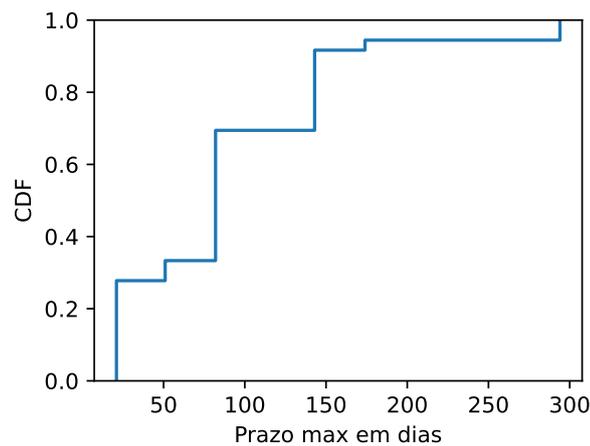


Figura 11 – Distribuição acumulada geral das atividades.

Outra maneira de visualizar as atividades pendentes é por meio da criação do CDF de todas as atividades pendentes, no qual é possível identificar visualmente os momentos que exigem maior atenção, pois são períodos em que as atividades se concentrarão, tornando necessário um planejamento mais criterioso após a definição da estratégia a ser adotada. Na Figura 11 observamos que pouco menos de 30% das atividades estão concentradas do 21º ao 50º dia, aproximadamente 60% do total pendente está previsto entre o 50º e 150º dia e os 10% restantes da área com prazo a partir de 150 dias. Sendo necessária atenção especial nesse primeiro intervalo, pois 30% da área está concentrada em um período curto de 29 dias.

4.3 Método da Pontuação Ponderada

Na seção anterior, abordamos de forma individual alguns fatores importantes para o planejamento e acompanhamento das atividades, demonstrando a complexidade inerente à operação e destacando pontos cruciais para a realização de um planejamento robusto, confiável e, principalmente, exequível. Apresentamos os resultados obtidos e algumas possibilidades de usos resultantes da manipulação dos dados, com foco em gerar informações de fácil interpretação e análise.

O processo de produção de florestas exige um plano de adubação bem definido e orientado. Dessa forma, a sugestão de criação de um plano orientado a dados mostrou-se extremamente válida e necessária. Com a utilização do método da pontuação ponderada, descrito no Capítulo 3, avaliamos o desempenho do ranqueamento produzido em dois cenários de interesse. Esse processo inicial permitiu o desenvolvimento de um modelo de automatização e priorização de atividades baseado nos critérios pré-determinados. A partir dos resultados obtidos nos dois cenários propostos (Tabela 1), foi possível analisar o desempenho do modelo.

Tabela 4 – Atividades mais e menos prioritárias utilizando os critérios do Cenário 1.

Atividades mais prioritárias									
Dcr. Projeto	Talhão	Data de Plantio	Atividade	Área residual	Prazo max.	Acesso ao talhão	Quantidade	Dose/ha	Ranqueamento Agr.
Irmãos Ferreira	NESBSB00143P23-018	2023-12-01	Calcário	16.3	21 dias	Sem restrição	8	1000.0	0.9
Cume	NESBSB00128R23-003	2023-12-01	1ª dose	4.6	21 dias	Com restrição	8	600.0	0.9
Irmãos Ferreira	NESBSB00143P23-019	2023-12-01	Calcário	8.3	21 dias	Sem restrição	8	1000.0	0.9
Irmãos Ferreira	NESBSB00143P23-035	2023-12-01	Calcário	26.4	21 dias	Sem restrição	8	2000.0	0.9
Irmãos Ferreira	NESBSB00143P23-015	2023-12-01	Calcário	13.6	21 dias	Sem restrição	8	1500.0	0.9
Irmãos Ferreira	NESBSB00143P23-033	2023-12-01	Calcário	8.2	21 dias	Sem restrição	8	1500.0	0.8
Placa	NESBSB00163R23-086	2023-12-01	1ª dose	41.5	21 dias	Restrição parcial	4	1200.0	0.8
Cume	NESBSB00128P23-011	2023-12-01	Calcário	5.5	21 dias	Com restrição	8	1500.0	0.8
Irmãos Ferreira	NESBSB00143P23-027	2023-12-01	Calcário	14.6	21 dias	Sem restrição	8	2000.0	0.8
Irmãos Ferreira	NESBSB00143P23-036	2023-12-01	Calcário	8.3	21 dias	Sem restrição	8	2500.0	0.8
Atividades menos prioritárias									
Valéria I	NESBSB00172P24-022	2024-02-01	1ª dose	12.7	143 dias	Sem restrição	7	400.0	0.6
Carlos Hosken	NESBSB00120P24-073	2024-01-01	Calcário	21.0	51 dias	Sem restrição	1	2000.0	0.6
Valéria I	NESBSB00172P24-021	2024-02-01	1ª dose	7.6	143 dias	Sem restrição	7	500.0	0.6
Valéria I	NESBSB00172P24-023	2024-02-01	1ª dose	5.2	143 dias	Sem restrição	7	500.0	0.6
Valéria I	NESBSB00172P24-013	2024-02-01	1ª dose	4.1	143 dias	Sem restrição	7	500.0	0.6
Jararaca II	NESBSB00145P24-001	2024-05-01	1ª dose	36.4	143 dias	Sem restrição	1	500.0	0.5
Placa	NESBSB00163P24-083	2024-02-01	1ª dose	3.1	143 dias	Restrição parcial	4	500.0	0.5
Jararaca II	NESBSB00145P24-001	2024-05-01	Calcário	4.7	174 dias	Sem restrição	1	1500.0	0.3
Paraíso	NESBSB00159P23-002	2023-11-01	2ª dose	34.5	294 dias	Sem restrição	5	500.0	0.3
Paraíso	NESBSB00159P23-001	2023-11-01	2ª dose	20.8	294 dias	Sem restrição	5	500.0	0.2

Como explicado no Capítulo 3 e Tabela 1, os critérios do Cenário 1 adotam o peso 0 (zero) para o critério de “*Acesso ao talhão*”, pesos menores para os demais critérios e foca na priorização dos “*Talhões*” com “*Prazo máximo*”, próximo ao limite. Dessa forma, busca-se priorizar atividades mais influenciadas pelo “*Prazo máximo*”, seguidas das demais variáveis especificadas no modelo. As atividades mais e menos prioritárias são apresentadas na Tabela 4.

Essas tabelas demonstram que, enquanto o critério “*Prazo máximo*” exerce uma influência significativa sobre o ranqueamento das atividades mais prioritárias, as atividades menos prioritárias têm prazos mais longos e são menos afetadas por esse critério. A falta de um comportamento padronizado nos demais critérios sugere que esses fatores, embora relevantes, são secundários em relação ao prazo máximo, que é crítico para a priorização eficiente das atividades de adubação.

Movendo para o Cenário 2, em que o acesso ao “*Talhão*” se torna extremamente crítico, as atividades são priorizadas de maneira diferente. Neste cenário, o critério “*Acesso ao talhão*” recebe um peso significativo (Tabela 1), refletindo a necessidade de antecipar atividades em áreas de difícil acesso antes do período chuvoso. Isso visa evitar atrasos operacionais e garantir que as atividades possam ser concluídas dentro dos prazos estabelecidos, independentemente das condições climáticas adversas. Assim, o uso deste cenário deve ser alinhado às previsões climáticas para que sua aplicação seja eficaz.

Outro ponto relevante a ser avaliado é que, em condições normais, seguindo o prazo máximo, a sequência de adubação acompanha, de forma geral, a sequência de plantio. No entanto, com o ranqueamento utilizando mais critérios, os “*Talhões*” não seguem a mesma sequência, podendo ter “*Talhões*” de vários projetos diferentes na mesma sequência. Essa situação deve ser avaliada, pois pode comprometer a logística da operação e o aproveitamento da mão de obra. O critério “*Quantidade de talhões*” foi inserido para minimizar esse efeito, mas ainda é preciso atenção a essa condição.

Tabela 5 – Atividades mais e menos prioritárias utilizando os critérios do Cenário 2.

Atividades mais prioritárias									
Dcr. Projeto	Talhão	Data de Plantio	Atividade	Área residual	Prazo max.	Acesso ao talhão	Quantidade	Dose/ha	Ranqueamento Agr.
Cume	NESBSB00128R23-003	2023-12-01	1ª dose	4.6	21 dias	Com restrição	8	600.0	0.9
Florália	NESBSB00136P24-007	2024-02-01	1ª dose	35.2	143 dias	Com restrição	9	500.0	0.8
Cume	NESBSB00128P23-011	2023-12-01	Calcário	5.5	21 dias	Com restrição	8	1500.0	0.8
Florália	NESBSB00136P24-058	2024-03-01	1ª dose	27.1	143 dias	Com restrição	9	500.0	0.8
Florália	NESBSB00136R24-033	2024-02-01	Calcário	6.8	82 dias	Com restrição	9	1000.0	0.8
Placa	NESBSB00163R23-086	2023-12-01	1ª dose	41.5	21 dias	Restrição parcial	4	1200.0	0.7
Irmãos Ferreira	NESBSB00143P23-018	2023-12-01	Calcário	16.3	21 dias	Sem restrição	8	1000.0	0.6
Placa	NESBSB00163P24-083	2024-02-01	Calcário	19.9	82 dias	Restrição parcial	4	1000.0	0.6
Placa	NESBSB00163P24-078	2024-01-01	Calcário	22.0	51 dias	Restrição parcial	4	2000.0	0.6
Irmãos Ferreira	NESBSB00143P23-019	2023-12-01	Calcário	8.3	21 dias	Sem restrição	8	1000.0	0.6
Atividades menos prioritárias									
Cascapau	NESBSB00122R24-096	2024-02-01	Calcário	8.1	82 dias	Sem restrição	5	1000.0	0.4
Valéria I	NESBSB00172P24-023	2024-02-01	1ª dose	5.2	143 dias	Sem restrição	7	500.0	0.4
Valéria I	NESBSB00172P24-013	2024-02-01	1ª dose	4.1	143 dias	Sem restrição	7	500.0	0.4
Cascapau	NESBSB00122R24-088	2024-02-01	Calcário	3.5	82 dias	Sem restrição	5	1000.0	0.4
Valéria I	NESBSB00172R24-020	2024-02-01	Calcário	1.7	82 dias	Sem restrição	7	2000.0	0.4
Jararaca II	NESBSB00145P24-001	2024-05-01	1ª dose	36.4	143 dias	Sem restrição	1	500.0	0.3
Carlos Hosken	NESBSB00120P24-073	2024-01-01	Calcário	21.0	51 dias	Sem restrição	1	2000.0	0.3
Paraíso	NESBSB00159P23-002	2023-11-01	2ª dose	34.5	294 dias	Sem restrição	5	500.0	0.3
Paraíso	NESBSB00159P23-001	2023-11-01	2ª dose	20.8	294 dias	Sem restrição	5	500.0	0.2
Jararaca II	NESBSB00145P24-001	2024-05-01	Calcário	4.7	174 dias	Sem restrição	1	1500.0	0.2

Na Tabela 5, observamos um comportamento diferente ao apresentado pelo Cenário 1. Conforme desejado, com a inclusão do critério “*Acesso ao talhão*”, é possível observar uma tendência de priorizar “*Talhões*” com restrição de acesso e aqueles com prazos mais curtos. As atividades mais e menos prioritárias são apresentadas na Tabela 5. As atividades menos prioritárias são aquelas sem restrição de acesso e com prazos mais longos.

Os resultados das atividades de maior e menor prioridade são apresentados na Tabela 5. A priorização pelo valor gerado para o ranqueamento agregado pode resultar em “*Talhões*” com o mesmo valor de ranqueamento, mas devido a critérios diferentes. Como observado na linha 4 e 5 da Tabela 5, ambos os “*Talhões*” têm o mesmo valor de ranqueamento agregado, mas por critérios distintos. O “*Talhão*” “*NESBSB00136P24-058*” teve um alto valor de ranqueamento devido ao critério “*Dose/ha*”, enquanto o “*NESBSB00136R24-033*”, com um prazo maior, foi influenciado pelos critérios “*Área residual*”.

Apesar dos pontos de atenção levantados, a realização de ajustes constantes será essencial para manter o alinhamento dos resultados do modelo com as diretrizes e premissas da CENIBRA. O ranqueamento utilizando o método da aplicação ponderada visando priorizar as demandas conforme critérios pré-estabelecidos apresentou um resultado satisfatório para a priorização das atividades e tem um grande potencial para ser explorado na operação de silvicultura. A flexibilidade da metodologia permite adequações e refinamentos que podem melhorar ainda mais a eficiência e a eficácia do planejamento e execução das atividades de adubação, garantindo que os recursos sejam alocados de maneira mais precisa e que os prazos sejam cumpridos de forma mais confiável.

5 Considerações Finais

Neste trabalho, abordamos o problema da integração de dados e da tomada de decisões em larga escala no contexto da silvicultura, com foco específico na CENIBRA, uma empresa multinacional localizada em Minas Gerais, Brasil. A gestão eficiente das atividades de adubação em um ambiente tão complexo e diversificado como o da CENIBRA apresenta desafios significativos, incluindo a necessidade de integrar dados de múltiplas fontes, garantir a precisão das informações e tomar decisões informadas e ágeis.

A integração dos dados provenientes de diferentes planilhas e sistemas da CENIBRA foi uma etapa crucial desse estudo. Utilizamos a linguagem de programação Python para unificar essas bases de dados, criando uma estrutura coesa que permitiu a realização de análises mais aprofundadas. A análise exploratória dos dados (AED) revelou padrões importantes e identificou variáveis críticas que impactam diretamente o processo de adubação. Essa abordagem não apenas facilitou a compreensão dos dados, mas também forneceu informações valiosas que orientaram a definição dos critérios de priorização.

Implementamos o método da pontuação ponderada para ranquear as atividades de adubação, priorizando-as com base em critérios como prazo máximo, acesso ao talhão, quantidade de área residual e dose por hectare. A aplicação desse método em dois cenários distintos demonstrou sua eficácia em fornecer uma visão estruturada e quantitativa das atividades mais urgentes, permitindo uma alocação mais eficiente de recursos. Os resultados mostraram que o critério de prazo máximo tem um impacto significativo no ranqueamento das atividades, enquanto o acesso ao talhão se tornou crucial em condições de difícil acesso, especialmente antes do período chuvoso. Apesar do sucesso do método da pontuação ponderada, há oportunidades para melhorar e expandir esta abordagem. Trabalhos futuros poderiam explorar a aplicação de outros métodos de ranqueamento, como o *Borda Count*, aplicável em contextos onde múltiplos critérios devem ser considerados simultaneamente. Este método pode proporcionar uma visão mais equilibrada das prioridades, levando em conta a classificação relativa de cada critério.

Além disso, é essencial realizar uma avaliação quantitativa e qualitativa do modelo proposto em outros cenários da empresa, incluindo filiais e diferentes áreas de operação. A implementação de novos critérios e a adaptação do modelo a diferentes contextos operacionais podem aumentar sua robustez e aplicabilidade. Ferramentas de visualização de dados e relatórios automatizados devem continuar sendo desenvolvidas para facilitar a interpretação e a tomada de decisão pelos gestores.

Em resumo, a integração de dados e a aplicação de métodos avançados de análise demonstraram grande potencial para aprimorar a gestão das atividades de adubação na CENIBRA. A utilização de técnicas de análise de dados, como a AED, combinada com métodos de ranqueamento baseados em dados, trouxe mais confiabilidade, segurança e controle ao processo. A continuidade desta abordagem, com a exploração de novos métodos e a aplicação em diferentes cenários, promete contribuir significativamente para a eficiência e a sustentabilidade das operações florestais da CENIBRA.

Referências

- ARMSTRONG, L. *et al.* **Improving data management and decision support systems in agriculture.** [S.l.]: Burleigh Dodds Science Publishing Limited, 2020.
- ASHTON, M. S.; KELTY, M. J. **The practice of silviculture: applied forest ecology.** [S.l.]: John Wiley & Sons, 2018.
- BEHRENS, J. T. Principles and procedures of exploratory data analysis. **Psychological methods**, American Psychological Association, v. 2, n. 2, p. 131, 1997.
- BHATLA, S. C.; LAL, M. A.; KATHPALIA, R.; BHATLA, S. C. Plant mineral nutrition. **Plant physiology, development and metabolism**, Springer, p. 37–81, 2018.
- BLAGOJEVIĆ, B.; JONSSON, R.; BJÖRHEDEN, R.; NORDSTRÖM, E.-M.; LINDROOS, O. Multi-criteria decision analysis (mcda) in forest operations—an introductory review. **Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering**, Faculty of Forestry and Wood Technology, University of Zagreb, v. 40, n. 1, p. 191–2015, 2019.
- BRAGA, A. A gestão da informação. **Millenium**, Instituto Politécnico de Viseu, 2000.
- BRITO, J. O.; FERRAZ, E. S.; BARRICHELO, L.; COUTO, H. d. A adubação mineral e seus efeitos sobre os anéis de crescimento da madeira de pinus caribaea var. bahamensis. **IPEF**, v. 32, n. 1, p. 5–17, 1986.
- BROWN, P.; HESKETH, A.; WILLIAMS, S. **The mismanagement of talent: Employability and jobs in the knowledge economy.** [S.l.]: Oxford University Press, USA, 2004.
- BRUCE, P.; BRUCE, A.; GEDECK, P. **Practical statistics for data scientists: 50+ essential concepts using R and Python.** [S.l.]: O’Reilly Media, 2020.
- CHAUHAN, S.; SINGH, M.; AGGARWAL, A. K. Data science and data analytics: artificial intelligence and machine learning integrated based approach. **Data science and data analytics: opportunities and challenges**, CRC Press Boca Raton, FL, USA, v. 1, 2021.
- DATA, M. C.; KOMOROWSKI, M.; MARSHALL, D. C.; SALCICCIOLI, J. D.; CRUTAIN, Y. Exploratory data analysis. **Secondary analysis of electronic health records**, Springer, p. 185–203, 2016.
- DIABY, V.; CAMPBELL, K.; GOEREE, R. Multi-criteria decision analysis (mcda) in health care: a bibliometric analysis. **Operations research for health care**, Elsevier, v. 2, n. 1-2, p. 20–24, 2013.
- EVANS, J. Sustainable silviculture and management. **Planted forests: Uses, impacts and sustainability**, CABI Wallingford UK, p. 113–140, 2009.
- FERREIRA, C. A.; SILVA, H. D. da; FERREIRA, C. A.; SILVA, H. D. D. *et al.* **Formação de povoamentos florestais.** [S.l.]: Colombo: Embrapa Florestas, 2008., 2008.

FERREIRA, M. F. C. Chemometrics iii-revisiting the exploratory analysis of multivariate data. **QUIMICA NOVA**, SOC BRASILEIRA QUIMICA CAIXA POSTAL 26037, 05599-970 SAO PAULO, BRAZIL, v. 45, n. 10, p. 1251–1264, 2022.

FLINT, R. W.; FLINT, R. W. Basics of sustainable development. **Practice of Sustainable Community Development: A Participatory Framework for Change**, Springer, p. 25–54, 2013.

FOELKEL, C. Minerais e nutrientes das árvores dos eucaliptos: Aspectos ambientais, fisiológicos, silviculturais e industriais acerca dos elementos inorgânicos presentes nas árvores. **Eucalyptus newsletter**, v. 2, 2005.

GIBBS, M. R.; SHANKS, G.; LEDERMAN, R. Data quality, database fragmentation and information privacy. **Surveillance & Society**, v. 3, n. 1, 2005.

GONÇALVES, J. *et al.* Recomendações de adubação para eucalyptus, pinus e espécies nativas. **Instituto de Pesquisa e Estudos Florestais, Piracicaba, SP. Disponível em:** < <http://www.ipef.br/silvicultura/adubacao.asp>>. Acesso em, v. 18, 2015.

GREBNER, D. L.; BETTINGER, P.; SIRY, J. P.; BOSTON, K. **Introduction to forestry and natural resources**. [S.l.]: Academic press, 2021.

HAN, J.; PEI, J.; TONG, H. **Data mining: concepts and techniques**. [S.l.]: Morgan kaufmann, 2022.

Irene Gomes and Umberlândia Cabral. **84,9% das indústrias de médio e grande porte utilizaram tecnologia digital avançada**. 2023. PINTEC Semestral, Editoria: Estatísticas Econômicas | Arte: Licia Rubinstein. Disponível em: <<https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/37973-84-9-das-industrias-de-medio-e-grande-porte-utilizaram-tecnologia-digital-avancada#:~:text=Todo%20o%20universo%20de%209.586,diferentes%20%C3%A1reas%2Ffun%C3%A7%C3%B5es%20de%20neg%C3%B3cios.>>

JENKINS, M.; SCHAAP, B. Forest ecosystem services. **Background analytical study**, v. 1, 2018.

JIawei, H.; MICHELINE, K. **Data mining: concepts and techniques**. [S.l.]: Morgan kaufmann, 2006.

JOSHI, A.; KAUSHIK, V. Big data and its analytics in agriculture. **Bioinformatics for agriculture: High-throughput approaches**, Springer, p. 71–83, 2021.

LATTERINI, F.; STEFANONI, W.; VENANZI, R.; TOCCI, D.; PICCHIO, R. Gis-ahp approach in forest logging planning to apply sustainable forest operations. **Forests**, MDPI, v. 13, n. 3, p. 484, 2022.

LIMA, L. P. de; FERREIRA, A. G.; VAZ, L. R. L.; ANDRADE, W. R.; FERREIRA, C. d. F. S. Nota técnica: Método para seleção de sistemas de tratamento de efluentes em agroindústrias de micro e pequeno porte. **Revista Engenharia na Agricultura**, Revista Engenharia na Agricultura, v. 26, n. 4, p. 334–342, 2018.

LOPES, A. S. **Manual de fertilidade do solo**. [S.l.]: Anda, 1989.

- MASCHI, L. F. C. Sumarização de dados no nodo por parâmetros: fusão de dados local em ambiente internet das coisas. Universidade Estadual Paulista (Unesp), 2018.
- MCKINNEY, W. **Python for data analysis**. [S.l.]: "O'Reilly Media, Inc.", 2022.
- MOREIRA, J.; SIMIONI, F. J.; OLIVEIRA, E. B. de. Importância e desempenho das florestas plantadas no contexto do agronegócio brasileiro. *Floresta*, Curitiba, v. 47, n. 1, p. 85-94, jan./mar. 2017., 2017.
- MORETTIN, P. A.; SINGER, J. d. M. Estatística e ciência de dados. 2022.
- MYATT, G. J. **Making sense of data: a practical guide to exploratory data analysis and data mining**. [S.l.]: John Wiley & Sons, 2007.
- PATIL, S. Exploratory data analysis shashankgoud patil1, nagaraja gs 2. 2020.
- PAWSON, S. M.; BRIN, A.; BROCKERHOFF, E. G.; LAMB, D.; PAYN, T. W.; PAQUETTE, A.; PARROTTA, J. Plantation forests, climate change and biodiversity. **Biodiversity and Conservation**, Springer, v. 22, p. 1203–1227, 2013.
- PÉREZ, J. P. Efeito do burnout e a sobrecarga na qualidade de vida no trabalho. **Estudios gerenciales**, v. 29, n. 129, p. 445–455, 2013.
- PERRY, D. A. The scientific basis of forestry. **Annual Review of Ecology and systematics**, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 29, n. 1, p. 435–466, 1998.
- PESODE, P.; BARVE, S.; WANKHEDE, S. V.; JADHAV, D. R.; PAWAR, S. K. Titanium alloy selection for biomedical application using weighted sum model methodology. **Materials Today: Proceedings**, Elsevier, v. 72, p. 724–728, 2023.
- RAIJ, B. van. **Recomendações de adubação e calagem para o Estado de São Paulo**. [S.l.]: IAC Campinas, 1997. v. 2.
- RAMANTSWANA, M.; GUERRA, S. P. S.; ERSSON, B. T. Advances in the mechanization of regenerating plantation forests: A review. **Current forestry reports**, Springer, v. 6, p. 143–158, 2020.
- REETZ, H. F. Fertilizantes e o seu uso eficiente. **São Paulo: ANDA**, v. 178, 2017.
- RIBEIRO, F.; FERREIRA, P.; ARAÚJO, M. Evaluating future scenarios for the power generation sector using a multi-criteria decision analysis (mcda) tool: The portuguese case. **Energy**, Elsevier, v. 52, p. 126–136, 2013.
- RODRIGUES, F. L. Metaheurística e sistema de suporte à decisão no gerenciamento de recursos florestais. Universidade Federal de Viçosa, 2001.
- ROY, S.; SHARMA, P.; NATH, K.; BHATTACHARYYA, D. K.; KALITA, J. K. Pre-processing: a data preparation step. **Encyclop Bioinform Comput Biol ABC Bioinform**, v. 463, 2018.
- SAHOO, K.; SAMAL, A. K.; PRAMANIK, J.; PANI, S. K. Exploratory data analysis using python. **International Journal of Innovative Technology and Exploring Engineering**, v. 8, n. 12, p. 4727–4735, 2019.

SHABAEV, A.; PITUKHIN, E. Integrated approach to automation of a wide range of forestry industry enterprises based on universal platform solutions. In: SPRINGER. **Data Science and Intelligent Systems: Proceedings of 5th Computational Methods in Systems and Software 2021, Vol. 2.** [S.l.], 2021. p. 870–883.

SILVA, G. R. da. Análise de mercado como ferramenta para o terceiro setor market analysis as a for the third sector.

SILVA, G. R. da. Análise de mercado como ferramenta para o terceiro setor. **Revista de Gestão, Finanças e Contabilidade**, v. 1, n. 1, p. 58–71, 2012.

SILVA, P. H. M. d. **Impactos das doses e do parcelamento da fertilização na produtividade, lixiviação e ciclagem de nutrientes em plantações de eucalipto.** Tese (Doutorado) — Universidade de São Paulo, 2011.

SINGH, V. Forest resources. In: **Textbook of Environment and Ecology.** [S.l.]: Springer, 2024. p. 143–153.

SLACK, N.; CHAMBERS, S.; JOHNSTON, R. *et al.* **Administração da produção.** [S.l.]: Atlas São Paulo, 2009. v. 2.

VENANZI, R.; LATTERINI, F.; CIVITARESE, V.; PICCHIO, R. Recent applications of smart technologies for monitoring the sustainability of forest operations. **Forests**, MDPI, v. 14, n. 7, p. 1503, 2023.

VETTORAZZI, C. A.; FERRAZ, S. F. d. B. Silvicultura de precisão: uma nova perspectiva para o gerenciamento de atividades florestais. **Agricultura de Precisão. Viçosa, MG: Editora da Universidade Federal de Viçosa**, p. 65–75, 2000.

WAGNER, M. B. Aspectos básicos da descrição e sumarização de informações em medicina. **Jornal de pediatria. Porto Alegre. V. 74, n. 1 (1998), p. 71-76**, 1998.

WATSON, H. J. Tutorial: Big data analytics: Concepts, technologies, and applications. **Communications of the Association for Information Systems**, v. 34, n. 1, p. 65, 2014.