

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**  
Escola de Engenharia  
Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos

Taís Resende Costa

**DESENVOLVIMENTO DE SENSORES VIRTUAIS PARA PREDIÇÃO DE CARBONATO  
RESIDUAL EM FORNOS DE CAL DA INDÚSTRIA DE CELULOSE**

Belo Horizonte  
2026

Taís Resende Costa

**DESENVOLVIMENTO DE SENSORES VIRTUAIS PARA PREDIÇÃO DE CARBONATO  
RESIDUAL EM FORNOS DE CAL DA INDÚSTRIA DE CELULOSE**

Dissertação apresentada ao Programa de Pós-graduação em Saneamento, Meio Ambiente e Recursos Hídricos da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Saneamento, Meio Ambiente e Recursos Hídricos.

Área de concentração: Meio Ambiente

Linha de pesquisa: Caracterização, prevenção e controle da poluição

Orientador: Marcelo Cardoso

Coorientador: Gustavo Matheus de Almeida

Belo Horizonte  
2026

C838d	<p>Costa, Taís Resende. Desenvolvimento de sensores virtuais para predição de carbonato residual em fornos de cal da indústria de celulose [recurso eletrônico] / Taís Resende Costa. - 2026. 1 recurso online (90 f. : il., color.) : pdf.</p> <p>Orientador: Marcelo Cardoso. Coorientador: Gustavo Matheus de Almeida.</p> <p>Dissertação (mestrado) Universidade Federal de Minas Gerais, Escola de Engenharia.</p> <p>Bibliografia: f. 80-90. Exigências do sistema: Adobe Acrobat Reader.</p> <p>1. Engenharia Sanitária - Teses. 2. - Aprendizado do computador - Teses. 3. Carbonatos - Teses. I. Cardoso, Marcelo. II. Almeida, Gustavo Matheus de. III. Universidade Federal de Minas Gerais. Escola de Engenharia. IV. Título.</p> <p style="text-align: right;">CDU: 628(043)</p>
-------	---



UNIVERSIDADE FEDERAL DE MINAS GERAIS  
Escola de Engenharia  
Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos

## FOLHA DE APROVAÇÃO

Desenvolvimento de Sensores Virtuais para Predição de Carbonato Residual em Fornos de Cal da Indústria de Celulose

**TAÍS RESENDE COSTA**

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

Prof. MARCELO CARDOSO  
GUSTAVO MATHEUS DE ALMEIDA  
Profa ELÉM PATRÍCIA ALVES ROCHA  
Profª SÍLVIA MARIA ALVES CORRÊA OLIVEIRA

Versão Final aprovada por:

Coordenação do PPG-SMARH

MARCELO CARDOSO  
Orientador

Belo Horizonte, 23 de fevereiro de 2026.



Documento assinado eletronicamente por **Gustavo Matheus de Almeida, Usuário Externo**, em 02/03/2026, às 09:51, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Marcelo Cardoso, Professor do Magistério Superior**, em 02/03/2026, às 12:15, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Sílvia Maria Alves Correa Oliveira, Professora do Magistério Superior**, em 02/03/2026, às 17:01, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Elem Patricia Alves Rocha, Usuária Externa**, em 04/03/2026, às 12:46, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Eduardo Coutinho de Paula, Coordenador(a) de curso de pós-graduação**, em 16/03/2026, às 11:36, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



A autenticidade deste documento pode ser conferida no site [https://sei.ufmg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **4988301** e o código CRC **8619FE7D**.

---

Referência: Processo nº 23072.209521/2026-20

SEI nº 4988301

*Aos meus pais, que transformaram  
esforço em oportunidade e amor em  
futuro.*

## **AGRADECIMENTOS**

Agradeço à Universidade Federal de Minas Gerais e seus docentes e servidores, cujo trabalho e dedicação foram fundamentais para minha formação. À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pelo apoio financeiro concedido.

Ao meu orientador, Marcelo Cardoso, por todas as conversas e ensinamentos, pelo apoio e atenção, por acreditar em mim, abrir portas e ir além do acadêmico, contribuindo de forma decisiva para minha formação profissional e pessoal.

Ao meu coorientador, Gustavo Almeida, pela paciência, orientação e apontamentos necessários para o desenvolvimento deste trabalho.

Aos professores Esa Vakkilainen e Falah Alobaid pela oportunidade de realizar o intercâmbio na LUT. À Cássia pelas dicas de sobrevivência na Finlândia e ao Jussi Saari por toda ajuda e pelas experiências finlandesas.

Ao grupo de pesquisa, por todas as trocas de conhecimento. Em especial, à Carla, por dividir comigo esta caminhada, e ao Diego, por estar sempre presente, pelo cuidado e por dividir desafios, conquistas e o cotidiano para além do acadêmico.

Aos meus pais, por acreditarem no poder transformador da educação e por todo o apoio e dedicação que tornaram possível minha chegada até aqui. À minha irmã, minha inspiração e meu porto seguro. A tia Elieste e Lucas por me acolherem mais uma vez em Belo Horizonte.

Por fim, agradeço à banca avaliadora, as professoras Sílvia e Elém, por se disponibilizarem a avaliar este trabalho, por todas as discussões e contribuições.

## RESUMO

No processo *Kraft*, o forno de cal converte  $\text{CaCO}_3$  em  $\text{CaO}$  para reutilização na caustificação, sendo o teor de carbonato residual um dos principais indicadores da eficiência da calcinação e da qualidade da cal produzida. Esse parâmetro é tradicionalmente determinado por meio de análises laboratoriais, o que introduz atrasos no controle do processo e limita a realização de ajustes operacionais em tempo hábil. O presente trabalho teve como objetivo desenvolver um modelo preditivo, baseado em dados de processo, para estimar o teor de carbonato residual em um forno de cal de uma indústria de celulose brasileira. Foram avaliados os algoritmos de aprendizado de máquina, redes neurais artificiais do tipo *Perceptron* Multicamadas (MLP) e modelos baseados em árvores de decisão, especificamente *Random Forest* e *XGBoost*. De forma geral, o segundo tipo apresentou desempenho superior aos modelos de redes neurais artificiais. Os resultados evidenciaram a importância de se considerar o tempo de residência do forno de cal, bem como a aplicação de técnicas adequadas de pré-processamento de dados. A suavização dos sinais, realizada com filtro de Savitzky–Golay, mostrou-se necessária e não comprometeu a preservação das tendências do processo. O melhor desempenho foi obtido com o algoritmo *XGBoost*, que permitiu estimar o teor de carbonato residual com frequência de trinta minutos a partir de variáveis de processo medidas online. O modelo apresentou coeficiente de correlação de 0,83 e erro absoluto médio de 0,33%. A análise de importância das variáveis indicou que a concentração de CO nos gases da chaminé, a temperatura do ar secundário e a velocidade do ventilador foram os parâmetros mais influentes na predição do carbonato residual. Foi realizada uma análise ambiental e econômica simplificada, que indicou um potencial de redução de emissões de  $\text{CO}_2$  de 0,0324  $\text{tCO}_2/\text{tCaO}$ , majoritariamente associada à economia de combustível, bem como um potencial de economia de energia térmica de 0,553  $\text{GJ/tCaO}$ , que correspondem, respectivamente, a 4% das emissões estequiométricas e 17% da energia gasta na produção de uma tonelada de  $\text{CaO}$ . Em termos econômicos, considerando a redução do consumo de gás natural a 18,956  $\text{US\$/GJ}$  e créditos de carbono a  $\text{US\$3,88}$ , o benefício econômico potencial associado ao controle da sobrequeima na unidade estudada poderia alcançar  $\text{US\$ 107.216,0}$  em 22 meses.

Palavras-chave: Carbonato residual; Modelagem preditiva; Aprendizado de máquina; Eficiência operacional; Sustentabilidade industrial.

## ABSTRACT

In the Kraft pulping process, the lime kiln is responsible for converting calcium carbonate ( $\text{CaCO}_3$ ) into calcium oxide ( $\text{CaO}$ ), which is reused in the causticizing plant. One of the main indicators of calcination efficiency and lime quality is the residual carbonate content, defined as the fraction of  $\text{CaCO}_3$  not converted into  $\text{CaO}$ . This parameter is traditionally determined through laboratory analyses, which introduce delays in process control and limit the ability to perform timely operational adjustments. The objective of this work was to develop a data-driven predictive model to estimate the residual carbonate content in a lime kiln of a Brazilian pulp mill. Different machine learning algorithms were evaluated, including Multilayer Perceptron (MLP) artificial neural networks, Random Forest, and XGBoost. Overall, tree-based models outperformed artificial neural network models. The results highlighted the importance of accounting for the lime kiln residence time, as well as the application of appropriate data preprocessing techniques. Signal smoothing using the Savitzky–Golay filter proved to be necessary and did not compromise the preservation of process trends. The best performance was achieved with the XGBoost algorithm, which enabled the estimation of residual carbonate content at a thirty-minute frequency using online process variables. The model achieved a correlation coefficient of 0,83 and a mean absolute error of 0,33%. Variable importance analysis indicated that CO concentration in the stack flue gas, secondary air temperature, and fan speed were the most influential parameters in predicting residual carbonate content. In addition, a simplified environmental and economic analysis indicated a potential reduction in  $\text{CO}_2$  emissions of 0,0324  $\text{tCO}_2/\text{tCaO}$ , mainly associated with fuel savings, as well as a potential thermal energy savings of 0,553  $\text{GJ}/\text{tCaO}$ , which correspond, respectively, to 4% of stoichiometric emissions and 17% of the energy spent in the production of one ton of  $\text{CaO}$ . From an economic perspective, considering the reduced consumption of natural gas, priced at 18,956  $\text{US}/\text{GJ}$ , and carbon credits valued at  $\text{US}\$ 3,88$ , the benefit associated with controlling overburning at the studied unit could reach  $\text{US}\$ 107,216.00$  over the analyzed period of 22 months.

**Keywords:** Residual carbonate; Predictive modeling; Machine learning; Operational efficiency; Industrial Sustainability.

## LISTA DE FIGURAS

Figura 1: Fluxograma do processo <i>Kraft</i> de celulose. ....	19
Figura 2: Zonas do forno de cal da indústria de celulose. ....	21
Figura 3: Seção transversal de uma partícula de cal de 60mm.....	25
Figura 4: Esquema ilustrativo do processo de recarbonatação em condições de operação do forno abaixo de 800 °C.....	26
Figura 5: Frequência de publicação por ano relacionada a estudos sobre fornos de cal na indústria de celulose.....	36
Figura 6: Visão geral da pesquisa mundial sobre fornos de cal na indústria de celulose. ....	36
Figura 7: Acúmulo progressivo de publicações sobre fornos de cal na indústria de celulose por fonte ao longo do tempo. ....	38
Figura 8: Acúmulo progressivo de publicações associadas a tópicos-chave de pesquisa na área de forno de cal. ....	39
Figura 9: Rede de interações entre palavras-chave indexadas e de artigos.....	41
Figura 10: Amostra de cal coletada em um forno de cal da indústria de celulose.....	44
Figura 11: Pipeline de aprendizagem de máquina que guia a metodologia desta pesquisa.....	47
Figura 12: Esquema de partição de dados da validação cruzada aninhada. ....	52
Figura 13: Dispersão entre Energia Específica e Carbonato Residual com dados brutos. ....	61
Figura 14: Dados suavizados (recorte) - (a)Janela de 5 pontos. (b)Janela de 11 pontos. ....	62
Figura 15: Distribuição estratificada de Carbonato Residual nos conjuntos de treinamento/validação e teste.....	63
Figura 16: Resultados do modelo MLP. (a)Valores reais vs preditos. (b) Histograma de resíduos. (c) Distribuição de resíduos em função de valores preditos. ....	65
Figura 17: Resultados do modelo <i>Random Forest</i> . (a)Valores reais vs preditos. (b) Histograma de resíduos. (c) Distribuição de resíduos em função de valores preditos. ....	66
Figura 18: Resultados do modelo <i>XGBoost</i> . (a)Valores reais vs preditos. (b) Histograma de resíduos. (c) Distribuição de resíduos em função de valores preditos. ....	67

Figura 19: Validação do modelo <i>XGBoost</i> em função do número de árvores. ....	68
Figura 20: Resultados do modelo preditivo final. (a)Valores reais vs preditos. (b) Histograma de resíduos. (c) Distribuição de resíduos em função de valores preditos. ....	69
Figura 21: Ajuste do modelo para o carbonato residual, mostrando os valores observados, a série predita com correção de viés e as faixas de incerteza associadas às estimativas.....	70
Figura 22: Análise de importância das <i>features</i> . ....	71

## LISTA DE TABELAS

Tabela 1: Valores testados para os hiperparâmetros de cada modelo. ....	51
Tabela 2: Variáveis selecionadas como entrada para o modelo. ....	60
Tabela 3: Resultados do desempenho dos modelos no subconjunto de teste na abordagem pontual. ....	64
Tabela 4: Resultados do desempenho dos modelos no subconjunto de teste na abordagem considerando o tempo de residência.....	64

## LISTA DE ABREVIATURAS E SIGLAS

GEE – Gases de efeito estufa  
CaCO<sub>3</sub> – Carbonato de cálcio  
CaO – Óxido de cálcio  
CO<sub>2</sub> – Dióxido de carbono  
NaOH – Hidróxido de sódio  
Na<sub>2</sub>S – Sulfeto de sódio  
Na<sub>2</sub>CO<sub>3</sub> – Carbonato de sódio  
Ca(OH)<sub>2</sub> – Hidróxido de cálcio  
TRS – Total reduced sulfur  
SO<sub>2</sub> – Dióxido de enxofre  
NO<sub>x</sub> - Óxidos de nitrogênio  
CTO – Crude Tall Oil  
NPEs – Non-process elements  
GNC – Gases não condensáveis  
IQR – Interquartile Range  
IsoForest – Isolation Forest  
LOF – Local outlier factor  
OC-SVM – One-class support vector machines  
ABC – Artificial Bee Colony  
MARS – Multivariate Adaptive Regression Splines  
MLP – Multilayer Perceptron  
XGBoost – Extreme Gradient Boost  
RF – Random Forest  
IPMA – Indian Paper Manufacturers Association  
LUT – Lappeenranta-Lahti University of Technology  
U of T – University of Toronto  
TAPPI - Associação Técnica da Indústria de Celulose e Papel  
IPPTA - Indian Pulp & Paper Technical Association  
CFD - Fluidodinâmica computacional  
ETEs - Estações de Tratamento de Efluentes  
LMD – Lime Mud Dryer  
PIMS - Plant Information Management System

CO – Monóxido de Carbono  
RNA – Rede Neural Artificial  
RMSE - Root Mean Squared Error  
MAE – Mean Absolute Error  
r – coeficiente de correlação

## SUMÁRIO

1	INTRODUÇÃO .....	15
2	OBJETIVOS .....	18
2.1	Objetivo Geral .....	18
2.2	Objetivos Específicos .....	18
3	REVISÃO DA LITERATURA .....	19
3.1	O processo <i>Kraft</i> .....	19
3.1.1	Ciclo de recuperação .....	20
3.2	Forno de cal .....	21
3.2.1	Combustíveis .....	23
3.2.2	Carbonato residual .....	24
3.2.3	Modelagem e Controle .....	27
3.3	Revisão bibliométrica .....	34
4	VISITAS TÉCNICAS .....	43
5	METODOLOGIA .....	45
5.1	Estudo de caso .....	45
5.2	Pré-processamento .....	48
5.3	Desenvolvimento do modelo .....	49
5.4	Pós-processamento .....	53
5.5	Análise de importância das <i>features</i> .....	54
5.6	Análise ambiental e econômica .....	55
6	RESULTADOS E DISCUSSÃO .....	59
6.1	Estudo de caso .....	59
6.2	Pré-processamento .....	60
6.3	Desenvolvimento do modelo .....	63
6.4	Pós-processamento .....	70
6.5	Análise de importância das <i>features</i> .....	71
6.6	Análise ambiental e econômica .....	72
7	CONCLUSÕES .....	75
8	RECOMENDAÇÕES .....	78
	REFERÊNCIAS .....	80

## 1 INTRODUÇÃO

A celulose é uma matéria-prima fundamental para a fabricação de uma ampla variedade de produtos, como papel, itens de higiene pessoal, embalagens, tecidos, entre outros (Argyropoulos *et al.*, 2023). O setor de celulose e papel possui relevância estratégica para a economia, tanto nacional quanto internacional, destacando-se pelo alto nível de desenvolvimento tecnológico, pela operação de unidades industriais de grande escala e por uma extensa base de florestas plantadas. O Brasil ocupa posição de destaque nesse cenário, sendo o segundo maior produtor mundial de celulose e o maior exportador do planeta. Em 2024, a produção nacional atingiu 25,5 milhões de toneladas, enquanto as exportações totalizaram 18,6 milhões de toneladas (Ibá, 2025).

Além de sua relevância econômica, a indústria de celulose ocupa posição de destaque na matriz energética, desempenhando papel significativo na dinâmica de geração e consumo de energia no setor industrial. Globalmente, é responsável por cerca de 5% da demanda energética da indústria, percentual que sobe para aproximadamente 16% no contexto industrial brasileiro (EPE, IEA e IBÁ; 2022). Ademais, destaca-se como a principal produtora e consumidora de energia renovável e biocombustíveis entre todos os segmentos industriais (Suárez *et al.*, 2018). Segundo a Empresa de Pesquisa Energética, em 2023 o licor preto, diretamente associado à indústria de celulose, contribuiu com 3,4% da Oferta Interna de Energia do Brasil, que é toda a energia necessária para movimentar a economia do país (EPE, 2025).

Assim como as demais áreas de manufatura, o setor também tem enfrentado diversos desafios relacionados à sustentabilidade, a mudanças climáticas e aspectos ambientais (Viitala *et al.*, 2025). Questões relacionadas ao uso da água, manejo florestal, biodiversidade e geração de resíduos têm recebido atenção cada vez maior nos relatórios corporativos. É uma indústria que apresenta uso intensivo de energia, água e madeira, além de gerar quantidades significativas de efluentes líquidos e resíduos sólidos (Sousa *et al.*, 2023). Atualmente, o setor de papel e celulose é responsável por aproximadamente 1,3% das emissões globais de gases de efeito estufa (GEEs) (Saira e Vauhkonen, 2025).

Por outro lado, o processo *kraft*, polpação química alcalina da madeira, principal rota de produção de celulose, tem potencial para reduzir seu próprio consumo de energia e contribuir para a descarbonização de outros setores (Rogerson *et al.*, 2024). Destaca-se nesse processo o ciclo de recuperação química, responsável por regenerar os insumos utilizados no processo de polpação e minimizar impactos ambientais (Mobarakeh, Silva e Kienberger, 2021). Esse ciclo favorece a sustentabilidade econômica e ecológica das fábricas, uma vez que permite a recuperação de reagentes químicos, a redução de resíduos e a geração de energia renovável no próprio processo (Anupam *et al.*, 2025).

Dentre as etapas deste ciclo, o forno de cal é uma unidade central da etapa de caustificação e calcinação, responsável pela transformação da lama de cal, essencialmente carbonato de cálcio ( $\text{CaCO}_3$ ), em óxido de cálcio ( $\text{CaO}$ ). Sua estabilidade operacional influencia diretamente a qualidade da cal produzida, o consumo de energia térmica e a eficiência global do processo *kraft* (Dylong e Rahman, 2024). Um dos principais indicadores de qualidade do processo é o carbonato residual, ou seja, a porcentagem de  $\text{CaCO}_3$  que não foi reagida. No entanto, apesar de sua importância, a sua determinação em fábricas de celulose é geralmente feita por análise laboratorial, tipicamente a cada quatro, seis ou oito horas (Ribeiro, 1998; Tran *et al.*, 2007). Essa baixa frequência de medição limita a capacidade de intervenção rápida e dificulta o controle em tempo real do processo. Nas poucas fábricas brasileiras em que já estão instalados sistemas de predição de carbonato residual, essas ferramentas ainda estão em fase de teste e experimentação.

Esse desafio pode ser superado por meio do desenvolvimento de modelos preditivos, com base em medições *on-line* de variáveis operacionais, para estimar, em tempo real, o teor de carbonato residual. Esses modelos orientados por dados, conhecidos como sensores virtuais, diferem dos sensores físicos convencionais que medem diretamente parâmetros como fluxo, pressão ou temperatura. Na prática, um sensor virtual é uma solução alternativa para prever um parâmetro-chave do processo que é difícil ou até impossível de medir diretamente no processo industrial (Oikonomou *et al.*, 2025).

Diante desse cenário, técnicas de modelagem e predição tornam-se estratégicas para aprimorar o controle operacional do forno de cal. O monitoramento avançado da calcinação favorece maior eficiência térmica, reduz o consumo de combustíveis e a reposição de reagentes, além de mitigar emissões de dióxido de carbono (CO<sub>2</sub>) associadas à sobrequeima e ao consumo excessivo de combustíveis.

Assim, a predição do carbonato residual assume papel estratégico na busca por fábricas de celulose mais eficientes, autossuficientes energeticamente e alinhadas aos compromissos globais de descarbonização. Portanto, o presente trabalho propõe um modelo preditivo baseado em dados para estimar o teor residual de carbonato em um forno de cal de uma fábrica de celulose *kraft* no Brasil.

A presente dissertação é dividida em oito capítulos, a saber: no capítulo 2 são apresentados os objetivos, geral e específicos, deste trabalho, o capítulo 3 apresenta uma revisão sobre o processo de produção *kraft*, com foco no forno de cal e seu controle; o capítulo 4 apresenta um breve relatório sobre as visitas técnicas realizadas em quatro fábricas de celulose, o capítulo 5 descreve a metodologia utilizada para a construção do modelo preditivo, o capítulo 6 apresenta os resultados e as discussões feitas sobre eles e os capítulos 7 e 8 apresentam, respectivamente, as conclusões e sugestões.

## **2 OBJETIVOS**

### **2.1 Objetivo Geral**

O presente trabalho tem por objetivo desenvolver um sensor virtual para estimar o teor de carbonato residual em fornos de cal da indústria de celulose *kraft*, com vista ao aprimoramento do controle operacional e à redução de impactos ambientais.

### **2.2 Objetivos Específicos**

Os seguintes objetivos específicos foram definidos para que o objetivo geral seja atingido:

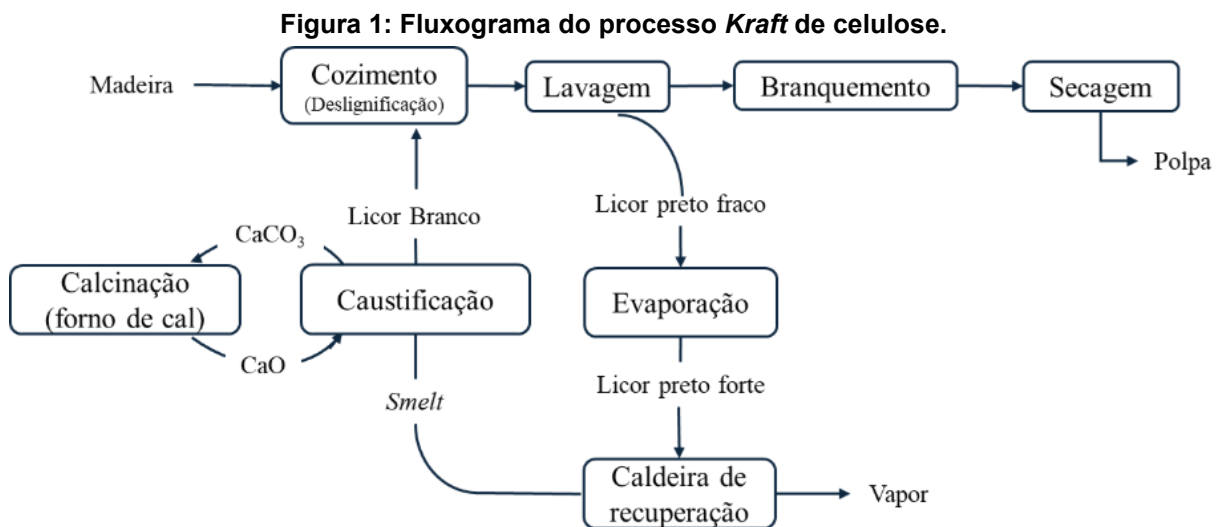
- Selecionar variáveis relevantes para serem utilizadas como entrada para o modelo de predição do carbonato residual, utilizando conhecimento de processo e critérios técnicos;
- Construir modelos preditivos candidatos para o carbonato residual, a partir de diferentes técnicas de aprendizagem de máquina;
- Avaliar a influência da suavização na distorção do sinal e na predição final em relação aos valores originais;
- Avaliar a relevância das variáveis selecionadas para o melhor modelo preditivo utilizando técnicas de análise de importância e correlacioná-las com o processo;
- Estimar o potencial de mitigação de emissões de CO<sub>2</sub> por redução de consumo de combustível e eliminação de situações de supercalcinação;
- Estimar simplificadamente a economia energética e financeira que poderia ser atingida com o uso do sensor virtual no período em que os dados foram coletados.

### 3 REVISÃO DA LITERATURA

Neste capítulo é apresentada a revisão da literatura pertinente à predição de carbonato residual, contemplando uma contextualização do problema, os principais conceitos, modelos e contribuições científicas que fundamentam a proposta deste estudo.

#### 3.1 O processo *Kraft*

O processo *Kraft* (Figura 1) é a principal rota de produção de celulose e corresponde a aproximadamente 90% da polpação química mundial (Argyropoulos *et al.*, 2023). Em 2023, mais de 150 milhões de toneladas de polpa celulósica foram produzidas desta forma no mundo (FAO, 2025). Dentre as vantagens em relação aos demais processos de polpação estão a preservação da cadeia de celulose, a possibilidade de uso de diversos tipos de madeira e uma alta eficiência de recuperação de compostos químicos (Tran e Vakkilainen, 2008). Ele pode ser dividido em duas grandes etapas, sendo elas a linha de fibras e o ciclo de recuperação.



Fonte: Autoria própria.

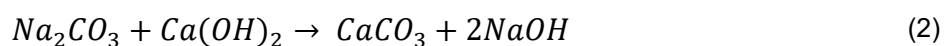
Inicialmente, os cavacos de madeira sofrem polpação em temperatura entre 150 e 170°C (Rahman, Avelin e Kyprianidis, 2020) e pressão de 700kPa (Neitzel e Kreczkuski, 2015) na presença de compostos alcalinos (principalmente, hidróxido de sódio, NaOH, e sulfeto de sódio, Na<sub>2</sub>S). Essa etapa resulta na separação das fibras de celulose, além de uma solução residual contendo compostos orgânicos e inorgânicos, conhecida como licor preto. A polpa celulósica segue na chamada linha

de fibras, um conjunto de etapas que inclui lavagem, branqueamento e secagem, resultando na formação das folhas (fardos) de celulose. Essa folha pode ser destinada à produção de papel em unidades fabris integradas ou ser expedida para processamento em outras instalações. O licor preto vai para o que é chamado de ciclo de recuperação, uma parte fundamental do processo que permite a geração de energia e o reaproveitamento dos compostos químicos utilizados no cozimento (Suárez *et al.*, 2018).

### 3.1.1 Ciclo de recuperação

O ciclo de recuperação é uma etapa decisiva tanto operacional quanto econômica e ambiental. Esta etapa engloba os processos de recuperação de compostos químicos utilizados no cozimento e de geração de vapor para recuperar energia a ser reutilizada no próprio processo. Dessa forma, contribui também para a diminuição da quantidade de efluente a ser tratada antes de ser descartada (Anupam *et al.*, 2025).

Primeiramente, o licor preto fraco que sai do digestor é concentrado nos evaporadores. O licor preto forte gerado é queimado na caldeira de recuperação e a combustão produz o *smelt*, uma solução fundida de compostos inorgânicos, que é então dissolvida em licor branco fraco para produzir licor verde (solução rica em carbonato de sódio,  $\text{Na}_2\text{CO}_3$ , e  $\text{Na}_2\text{S}$ ). Esta solução é clarificada e misturada com cal (óxido de cálcio,  $\text{CaO}$ ) nos apagadores e caustificadores. A cal então reage com a água da solução, conforme a Equação 1, gerando hidróxido de cálcio ( $\text{Ca(OH)}_2$ ). O hidróxido formado reage com o carbonato de sódio, conforme a Equação 2. O licor branco ( $\text{NaOH}$  e  $\text{Na}_2\text{S}$ ) é regenerado por meio da reação e é gerado um subproduto chamado lama de cal (essencialmente carbonato de cálcio,  $\text{CaCO}_3$ ). O forno de cal é responsável por converter essa lama de cal de volta em cal, por meio de um processo de calcinação em alta temperatura, tipicamente em torno de  $800\text{ }^\circ\text{C}$  (Equação 3), (Wu *et al.*, 2024). A cal gerada é então reutilizada na caustificação, fechando o ciclo (Assumpção *et al.*, 1988).



### 3.2 Forno de cal

O forno de cal é um equipamento essencial no processo de recuperação química da indústria de celulose, atuando como um reator e trocador de calor rotativo, no qual a energia térmica de uma fase gasosa quente é transferida para o material sólido do leito (Boateng, 2016) em um fluxo contracorrente. Trata-se de um cilindro de grandes proporções, levemente inclinado e revestido internamente com tijolos refratários, cuja rotação e inclinação favorecem o escoamento gradual da lama de cal da extremidade superior (alimentação) até a extremidade inferior, onde se localiza o queimador e ocorre a descarga do produto calcinado (Juneja *et al.*, 2020).

Segundo Francey, Tran e Berglin (2011), as dimensões médias dos fornos de cal utilizados na indústria de celulose eram de aproximadamente 92 metros de comprimento e 3,5 metros de diâmetro, com tendência de aumento da capacidade de produção (em toneladas de CaO por dia) ao longo dos anos. Em condições operacionais típicas, a rotação do forno varia entre 0,5 e 2 rotações por minuto (rpm), o que, em conjunto com a inclinação do corpo do forno, geralmente entre 1,5° e 3°, determina o tempo de residência da lama, que pode variar de 1,5 a 4 horas (Adams, 2008).

Quatro zonas funcionais são identificadas em um forno de cal (Figura 2) (Alves *et al.*, 2015). Na zona de secagem ocorre a evaporação da água presente na lama de cal, que entra no forno com um teor de aproximadamente 70% de sólidos. No segundo estágio ocorre o aquecimento da lama até a temperatura de calcinação. No terceiro estágio, já com o calor necessário, há a decomposição do  $\text{CaCO}_3$  em  $\text{CaO}$  e  $\text{CO}_2$ . Na zona de resfriamento, a cal recuperada é resfriada pelo ar secundário para descarga.



Fonte: ALVES *et al.*, 2015

Fornos de cal mais antigos eram tradicionalmente equipados com correntes na zona fria, em razão da menor temperatura dos gases nessa região, o que limita a eficiência da troca térmica entre os gases e a lama de cal. As correntes tinham como função principal ampliar a área superficial de contato, intensificando a transferência de calor e, conseqüentemente, favorecendo o processo de secagem da lama. Em contrapartida, fornos de cal mais recentes incorporam sistemas externos de secagem, usualmente compostos por ciclones e precipitadores eletrostáticos, que removem a umidade da lama antes de sua alimentação no forno. A adoção desses sistemas elimina a necessidade do uso de correntes internas, permitindo que toda a extensão do forno seja destinada predominantemente às etapas de aquecimento e calcinação, com ganhos em eficiência térmica e operacional (Adams, 2008; Manning, Hanson e Hart, 2017).

Apesar de ser uma operação central no ciclo de recuperação (Berglin, Von Schenck e Vakkilainen, 2024), o forno de cal é frequentemente citado por muitos autores como um gargalo no processo (Michniewicz e Janiga, 2010; Mateos-Espejel, Radiotis e Jeema, 2013; Argyropoulos *et al.*, 2023). Isso porque o processo de calcinação apresenta desafios operacionais significativos, decorrentes de sua dinâmica complexa e natureza multivariável. Entre os principais fatores que dificultam a operação eficiente estão a cinética de reação não linear, os longos atrasos de tempo inerentes ao processo térmico e as variações nas características da alimentação da lama de cal, que comprometem a estabilidade e o controle do sistema (Järvensivu, Saari e Jämsä-Jounela, 2001).

Dentre os problemas operacionais frequentemente relatados está a formação de anéis, que ocorre pelo acúmulo de material aderido à parede interna do forno, o que compromete sua eficiência térmica e estabilidade operacional (Ryan, Bussmann e DeMartini, 2022). Também é comum a formação de bolas (aglomerados esféricos de material), que gera danos ao revestimento refratário, levando à baixa eficiência energética e resultando em alto consumo de combustível e emissões atmosféricas relevantes, tais como material particulado, gases de enxofre reduzido (TRS), dióxido de enxofre (SO<sub>2</sub>) e óxidos de nitrogênio (NO<sub>x</sub>) (Tran e Vakkilainen, 2008).

Ainda do ponto de vista ambiental, em muitas fábricas de celulose o forno de cal é a única fonte de consumo de combustíveis fósseis, com exceção de pequenas quantidades utilizadas em outros equipamentos nas etapas de partida e parada da planta (Berglin, von Schenck & Vakkilainen, 2024). Portanto, esta etapa é a principal responsável pelas emissões de CO<sub>2</sub> fóssil do setor, enquanto as demais fontes são basicamente emissoras de CO<sub>2</sub> biogênico (Kuparinen, Vakkilainen e Tynjälä, 2019).

### 3.2.1 Combustíveis

Atualmente, a maior parte das plantas utilizam combustíveis fósseis como o óleo e o gás natural como principal fonte de energia no forno de cal (Kuparinen e Vakkilainen, 2017). No entanto, o setor tem cada vez mais investido em pesquisas para substituir os combustíveis fósseis por combustíveis renováveis devido aos custos e ao impacto ambiental das emissões atmosféricas (Suárez *et al.*, 2018). Esses combustíveis alternativos podem ser obtidos a partir do próprio processo de produção da polpa celulósica, como por exemplo metanol, hidrogênio e gases não condensáveis (Kuparinen e Vakkilainen, 2017). O metanol é separado dos condensados gerados na etapa de evaporação. O hidrogênio é produzido na planta química instalada junto ao complexo industrial, responsável pela fabricação de insumos utilizados no processo *kraft*. Os gases não condensáveis são subprodutos do próprio processo de produção. Embora historicamente tenham sido tratados como resíduos industriais, esses compostos hoje são valorizados e reintegrados ao ciclo produtivo, alinhando-se ao conceito de economia circular.

Berglin, von Schenck e Vakkilainen (2024) observaram um aumento no número de plantas industriais de celulose que passaram a empregar biocombustíveis como forma de substituir, ainda que parcialmente, combustíveis fósseis. Atualmente as fábricas operam com misturas de combustíveis, combinando fontes fósseis e renováveis para atender às demandas energéticas do processo de calcinação.

Estudos recentes (Taillon, Horvath e Oksman, 2018; Ahmadvand e Sowlati, 2023; Khadivi e Sowlati, 2024; Berglin, von Schenck & Vakkilainen, 2024) indicam que a gaseificação de biomassa para a produção de gás de síntese (*syngas*) tem ganhado destaque na indústria de celulose como uma alternativa promissora para a substituição parcial ou total de combustíveis fósseis nos fornos de cal. Essa tecnologia

permite o aproveitamento energético de resíduos lignocelulósicos e contribui para a redução das emissões de gases de efeito estufa, além de oferecer maior flexibilidade operacional.

Paralelamente, o uso do licor preto tem sido investigado como potencial fonte indireta de energia para os fornos, seja por meio da extração de lignina para combustão, seja como base para a produção de biocombustíveis ou gás de síntese (Tiller *et al.*, 2025). Além disso, Soto e Pereira (2017) relataram a aplicação de *crude tall oil* (CTO) como substituto parcial do óleo combustível em um forno de cal. O CTO é um coproduto do processo de cozimento *kraft* da madeira quando são utilizadas coníferas (pinus, produção de fibra longa), sendo resultado da extração de ácidos graxos e resinas presentes neste tipo de matéria-prima. Este estudo reforça a viabilidade técnica e econômica de alternativas renováveis obtidas na própria instalação.

No entanto, conforme relatado por Berglin, von Schenck e Vakkilainen (2024), o uso de combustíveis alternativos em fornos de cal exige cuidados operacionais e pesquisa aprofundada. Aspectos como a disponibilidade do combustível e os procedimentos de manutenção específicos associados à sua utilização devem ser cuidadosamente avaliados. Além disso, esses combustíveis impactam não apenas a operação do próprio forno, mas também o processo de caustificação e o equilíbrio do ciclo de recuperação química como um todo. Entre os efeitos observados estão a introdução de elementos não processuais (do inglês *non-process elements*, NPEs), alterações no perfil térmico do forno, na estabilidade da chama e nas emissões atmosféricas. Combustíveis como o metanol e gases não condensáveis (GNC) estão associados à emissão de TRS.

### 3.2.2 Carbonato residual

O carbonato residual corresponde à fração de  $\text{CaCO}_3$  que não foi convertida em  $\text{CaO}$  durante o processo de calcinação. A conversão ocorre a partir da superfície externa em direção ao interior das partículas, de modo que as partículas de cal apresentam, em geral, uma estrutura heterogênea, composta por uma camada externa rica em  $\text{CaO}$  e um núcleo interno constituído majoritariamente por  $\text{CaCO}_3$ . Esse comportamento pode ser observado na Figura 3, especialmente em partículas de maior diâmetro. Em contraste, partículas de menor tamanho, tipicamente inferiores a

20–30 mm, tendem a apresentar uma distribuição mais homogênea do carbonato residual ao longo de sua estrutura, em função da maior eficiência da transferência de calor e da difusão dos gases reativos (Tran *et al.*, 2007).

**Figura 3: Seção transversal de uma partícula de cal de 60mm.**



Fonte: Tran *et al.*, 2007

A medida de carbonato residual no produto do forno de cal é normalmente mantida entre 1,5% e 2,5%, faixa considerada ideal para equilibrar eficiência energética e qualidade da cal produzida, segundo a literatura (Tran, 2008). No entanto, na prática o carbonato residual é mantido entre 3 e 4,5% nas fábricas brasileiras. Esse controle é realizado, em grande parte, por meio do ajuste preciso da temperatura na zona de calcinação, a partir da vazão de combustível.

Temperaturas muito elevadas podem levar à supercalcinação ou sobrequeima, que pode degradar o revestimento refratário do forno, gerar perdas térmicas desnecessárias e produzir uma cal queimada, com um baixo teor de  $\text{CaCO}_3$  residual e menor eficiência no processo de caustificação. Por outro lado, temperaturas baixas podem levar à subcalcinação, que resulta em excesso de carbonato residual e novamente em cal com baixa reatividade (Tran, 2008).

Temperaturas de operação insuficientes (abaixo de  $800^\circ\text{C}$ ) também favorecem a ocorrência da recarbonatação, processo no qual a camada externa de  $\text{CaO}$  reage com o  $\text{CO}_2$  presente nos gases do forno, resultando na formação de  $\text{CaCO}_3$ . Esse

fenômeno contribui para o aumento do teor de carbonato residual no produto final, conforme ilustrado esquematicamente na Figura 4.

**Figura 4: Esquema ilustrativo do processo de recarbonatação em condições de operação do forno abaixo de 800 °C.**



Fonte: Tran, 2008.

O carbonato residual é, portanto, uma importante variável do processo de calcinação e também uma medida indireta de sua eficiência. Uma eficiência alta é essencial tanto do ponto de vista operacional quanto ambiental na indústria de celulose. Operacionalmente, uma calcinação eficiente assegura a conversão ideal do carbonato de cálcio em óxido de cálcio, reduzindo significativamente a necessidade de reposição externa de cal virgem (Alves *et al.*, 2015). Essa menor dependência de cal comercial representa uma economia relevante e traz benefícios ambientais expressivos, uma vez que reduz a demanda por extração de calcário em minas, diminuindo os impactos da mineração sobre o solo, a vegetação nativa e os recursos hídricos.

Além disso, a menor geração de resíduos sólidos associados ao descarte de lama não completamente calcinada contribui para um processo mais limpo e sustentável. No aspecto energético, a eficiência de calcinação também está relacionada ao uso otimizado de combustíveis (Guzzo, 2024), o que permite menor consumo e redução das emissões atmosféricas, incluindo gases de efeito estufa e compostos sulfurados. Com isso, promove-se não apenas a sustentabilidade do processo industrial, mas também o alinhamento com práticas mais responsáveis de gestão de recursos naturais e controle ambiental.

Apesar de ser uma variável crítica, na maior parte das empresas, atualmente o carbonato residual não é medido diretamente em tempo real (*on-line*). Sua quantificação geralmente é feita por análises laboratoriais (*off-line*), com coleta de amostras e procedimentos analíticos que demandam tempo. A frequência de análise varia de fábrica para fábrica, mas, de qualquer forma, traz uma limitação que dificulta o controle imediato do processo e a tomada de decisões.

### 3.2.3 Modelagem e Controle

Alguns estudos como o de Suárez e colaboradores (2018), Liu e Cheng (2024) e Othen e colaboradores (2025) constatam que a manufatura moderna é caracterizada pela geração intensiva de dados em todas as etapas do processo produtivo. Quando devidamente tratados e analisados, esses dados têm o potencial de gerar conhecimento profundo e relevante sobre os processos, contribuindo para a otimização, a melhoria contínua e a tomada de decisões mais assertivas em ambientes industriais complexos. Apesar disso, na indústria de celulose, alguns parâmetros importantes ainda não são monitorados de forma *on-line*. E, conforme destacado por Trung (2015), a falta de sensores *on-line* para importantes variáveis de operação dificulta o desenvolvimento de controle avançado de processo.

No entanto, trabalhar com medições reais traz um desafio adicional relacionado à complexidade dos processos industriais, que frequentemente apresentam comportamento não linear, forte acoplamento entre variáveis, dinâmicas multiescalares e influência de distúrbios não mensurados (Du *et al.*, 2024). Além disso, a qualidade dos dados industriais impõe limitações relevantes, uma vez que dados reais são afetados por ruído, valores ausentes e atípicos, mudanças de regime operacional e não estacionariedade (Ren *et al.*, 2025).

Neste contexto, o pré-processamento desempenha um papel crítico no desempenho e na confiabilidade de modelos preditivos, sendo um dos fatores mais determinantes para a qualidade das previsões obtidas. Sua ausência pode levar a um aumento do erro de previsão e redução da capacidade de generalização, especialmente em abordagens baseadas em aprendizado de máquina (Du *et al.*, 2024). Técnicas de detecção e tratamento de *outliers*, filtragem e suavização de ruídos, normalização, seleção e extração de características, bem como o alinhamento temporal de variáveis, são fundamentais para capturar adequadamente a dinâmica do processo e reduzir efeitos indesejados da não estacionariedade dos dados (Ge *et al.*, 2017).

*Outliers* surgem devido a falhas mecânicas, mudanças no comportamento do sistema, fraudes (como manipulação, roubo, falsificação, etc.), erro humano, erro de instrumento ou simplesmente por desvios naturais (Hodge e Austin, 2004). A remoção destes valores é fundamental, visto que observações anômalas podem distorcer o

ajuste dos modelos, comprometer a estabilidade numérica e reduzir a acurácia das previsões. Dentre os métodos estatísticos univariados, destacam-se o *z-score* e o intervalo interquartil (*Interquartile Range*, IQR).

O *z-score* baseia-se na padronização das observações em relação à média e ao desvio-padrão, assumindo implicitamente uma distribuição aproximadamente normal, de modo que valores com magnitude elevada indicam baixa probabilidade de ocorrência e, portanto, potencial anomalia (Hodge & Austin, 2004). O método IQR, por sua vez, é fundamentado em estatísticas de posição, utilizando a diferença entre o terceiro e o primeiro quartis para definir limites inferiores e superiores, sendo menos sensível à presença de valores extremos e mais adequado para dados assimétricos ou com distribuições não gaussianas (Tukey, 1977).

Tais métodos estatísticos tradicionais são amplamente utilizados devido à sua simplicidade e interpretabilidade, porém apresentam limitações importantes quando aplicados a dados industriais de alta dimensionalidade, não linearidade e múltiplos regimes operacionais (Hodge e Austin, 2004). Em contraste, métodos estatísticos e de aprendizado de máquina como *k-means*, *local outlier factor* (LOF) e *one-class support vector machines* (OC-SVM) tem ganhado destaque. Tais técnicas dispensam a adoção de hipóteses estatísticas prévias, identificando anomalias por meio da organização e distribuição dos dados.

Dentre estas abordagens, o *Isolation Forest* (IsoForest), em particular, define *outliers* como observações que podem ser isoladas com menor número de partições aleatórias no espaço de atributos, explorando o fato de que anomalias tendem a ocupar regiões esparsas do conjunto de dados, o que resulta em boa escalabilidade, robustez em alta dimensionalidade e independência de hipóteses distributivas (Liu *et al.*, 2008). Essas características tornam o IsoForest especialmente adequado para dados industriais complexos (Ge *et al.*, 2017).

Além da remoção de *outliers*, a suavização de dados também é uma etapa essencial no pré-processamento de sinais industriais, pois permite reduzir o impacto de ruídos de medição e flutuações de alta frequência que podem comprometer tanto a análise exploratória quanto o desempenho de modelos preditivos (Montgomery, 2024). Entre as técnicas mais empregadas, as médias móveis são amplamente utilizadas devido à

sua simplicidade e eficiência na atenuação de ruído aleatório em séries temporais, embora possam introduzir atraso temporal e perda de informação dinâmica quando janelas excessivamente longas são adotadas (Montgomery, 2024).

O filtro de *Hampel*, por sua vez, combina suavização com detecção de *outliers*, substituindo observações anômalas com base na mediana local e na dispersão dos dados, sendo particularmente adequado para sinais contaminados por picos espúrios (Hampel *et al.*, 1986). Já o filtro de *Savitzky–Golay* realiza a suavização por meio do ajuste local de polinômios via mínimos quadrados, permitindo reduzir o ruído preservando características importantes do sinal, como picos, curvaturas e derivadas, o que o torna especialmente adequado para aplicações nas quais a dinâmica local do processo deve ser mantida (Savitzky e Golay, 1964). Dessa forma, a escolha da técnica de suavização deve considerar as características do processo, o tipo de ruído presente e os objetivos da modelagem preditiva, uma vez que cada método envolve compromissos distintos entre redução de ruído, preservação da informação e custo computacional (Ge *et al.*, 2017).

Após a realização do pré-processamento adequado, o desenvolvimento do modelo preditivo também envolve uma série de escolhas e análises a serem feitas. A decisão final deve considerar o compromisso entre desempenho preditivo, interpretabilidade, custo computacional e requisitos operacionais da aplicação. Estudos demonstraram a capacidade de modelos baseados em aprendizagem de máquina funcionarem como modelos preditivos satisfatórios para ambientes industriais complexos, inclusive para a indústria de celulose e para fornos de cal. Isso porque têm capacidade de aprender relações complexas diretamente a partir dos dados, sem a necessidade de formulações explícitas baseadas em primeiros princípios (Russell & Norvig, 2021).

Lianhua e colaboradores (2011) investigaram a aplicação de um novo algoritmo de controle inferencial preditivo no processo de lavagem de polpa, evidenciando ganhos significativos na velocidade de resposta e na eficiência do monitoramento do processo. Huang e colaboradores (2017) aplicaram algoritmos de aprendizagem de máquina para otimizar a previsão da qualidade da polpa na indústria de papel e celulose. Belisário e colaboradores (2023) exploraram a aplicação de redes neurais artificiais para predição e análise de emissões de SO<sub>2</sub> e material particulado em

caldeiras de recuperação. Järvensivu, Saari e Jämsä-Jounela (2001) usaram modelos de redes neurais combinados com lógica fuzzy e equação linguística para construir um sistema inteligente de controle de forno. Suárez e colaboradores (2018) utilizaram uma técnica baseada na combinação do algoritmo *Artificial Bee Colony* (ABC) ou de colônia artificial de abelhas com regressões adaptativas por splines multivariadas (do inglês *Multivariate Adaptive Regression Splines*, MARS) para construir um modelo preditivo de consumo de gás natural e emissões de SO<sub>2</sub> e NO<sub>x</sub> em fornos de cal da indústria de celulose.

Em relação ao carbonato residual, Ribeiro (1988) reportou o uso de redes neurais *Perceptron* Multicamadas (MLP) para predição da disponibilidade da cal, em que o teor de carbonato residual era uma das variáveis de saída, a partir de dados simulados. Guzzo (2022) relata que um modelo também construído com MLP foi capaz de estimar de forma eficaz a eficiência de calcinação de um forno de uma indústria específica. O autor destaca ainda o potencial e a importância de algoritmos preditivos no contexto industrial.

Os trabalhos analisados evidenciam a relevância da modelagem de processos, do controle avançado, da predição de variáveis operacionais e da aplicação de modelos de aprendizagem de máquina na indústria de celulose. Em um setor caracterizado por processos complexos, altamente integrados e energeticamente intensivos, essas abordagens têm se mostrado fundamentais para a otimização do desempenho operacional, redução de custos, aumento da eficiência energética e minimização dos impactos ambientais, incluindo uma redução de emissões atmosféricas, que são relativamente menores em plantas operando com estabilidade. A literatura, portanto, aponta para uma tendência de incorporação dessas tecnologias como parte essencial da modernização e da sustentabilidade no setor. Entre as técnicas mais utilizadas nesse contexto, destacam-se os métodos baseados em árvores de decisão e as redes neurais artificiais, que diferem tanto em estrutura quanto em interpretabilidade e requisitos computacionais.

As redes neurais artificiais constituem uma classe de modelos inspirados no funcionamento do sistema nervoso biológico, sendo amplamente utilizadas em problemas de regressão e classificação devido à sua capacidade de modelar relações

altamente não lineares entre variáveis de entrada e saída, no caso de aprendizagem supervisionada (Russell e Norvig, 2021). Esses modelos, em geral, são estruturados em camadas de neurônios interconectados, nos quais cada neurônio realiza uma combinação linear das entradas seguida da aplicação de uma função de ativação não linear, permitindo a construção de representações hierárquicas dos dados. Uma propriedade fundamental das redes neurais é sua capacidade de aproximação universal, que garante que, sob determinadas condições, redes com pelo menos uma camada oculta podem aproximar qualquer função contínua com precisão arbitrária (Hornik, Stinchcombe e White, 1989). Em aplicações industriais, redes neurais são particularmente atrativas para processos complexos e multivariados, embora apresentem desafios relacionados à interpretabilidade, ao risco de sobreajuste e à necessidade de um pré-processamento cuidadoso dos dados (Ge *et al.*, 2017).

As redes neurais artificiais do tipo MLP pertencem à classe das redes *feedforward*, nas quais a informação flui unidirecionalmente da camada de entrada em direção à camada de saída, sem a presença de realimentação interna ou ciclos (Russell e Norvig, 2021). Essas redes são estruturadas em múltiplas camadas de neurônios, geralmente compostas por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Em cada neurônio, as entradas são combinadas linearmente por meio de pesos sinápticos, somadas a um termo de viés e, em seguida, transformadas por uma função de ativação não linear, como ReLU, sigmoide ou tangente hiperbólica. Essa não linearidade é fundamental para que a rede seja capaz de modelar relações complexas entre variáveis de processo.

O treinamento de uma MLP é realizado, em geral, por meio do algoritmo de retropropagação do erro (*backpropagation*), no qual os pesos da rede são ajustados iterativamente com base no gradiente de uma função de custo, utilizando métodos de otimização como descida do gradiente ou suas variantes (Rumelhart *et al.*, 1986). Nesse processo, o erro calculado na camada de saída é propagado de volta pelas camadas ocultas, permitindo atualizar os parâmetros da rede de forma a minimizar a discrepância entre as previsões do modelo e os valores observados. A capacidade das MLPs de aproximar funções não lineares arbitrárias é garantida teoricamente pelo teorema da aproximação universal, que estabelece que redes *feedforward* com ao

menos uma camada oculta podem aproximar qualquer função contínua sob condições adequadas (Hornik, Stinchcombe e White, 1989).

Em aplicações de modelagem preditiva de processos industriais, as MLPs destacam-se pela flexibilidade e pela capacidade de capturar interações complexas entre múltiplas variáveis de entrada. No entanto, essas redes apresentam particularidades importantes, como a sensibilidade à escala dos dados, o que torna essencial o uso de técnicas adequadas de normalização e pré-processamento. Além disso, o desempenho das MLPs depende fortemente da escolha da arquitetura da rede, incluindo o número de camadas ocultas, o número de neurônios por camada e as funções de ativação adotadas. A ausência de uma interpretação direta dos parâmetros do modelo e o risco de sobreajuste, especialmente em bases de dados limitadas, também representam desafios relevantes para sua aplicação em ambientes industriais críticos (Ge *et al.*, 2017; Russell e Norvig, 2021).

As árvores de decisão pertencem a uma classe de modelos preditivos baseados em regras, em que o espaço de atributos é particionado recursivamente por meio de testes condicionais, resultando em uma estrutura hierárquica de fácil interpretação (Russell e Norvig, 2021). Cada nó interno da árvore corresponde a uma decisão baseada em uma variável de entrada, enquanto os nós-folha fornecem a previsão do modelo. A principal vantagem das árvores de decisão reside em sua transparência e interpretabilidade, o que as torna especialmente atrativas em contextos industriais, onde a compreensão das relações entre variáveis é essencial. No entanto, árvores individuais tendem a apresentar alta variância e sensibilidade a pequenas perturbações nos dados, o que pode resultar em sobreajuste. Por essa razão, métodos de conjunto baseados em árvores, como *Random Forest* (RF) e *boosting*, são frequentemente empregados para melhorar a robustez e o desempenho preditivo, mantendo parte da interpretabilidade inerente à estrutura das árvores (Breiman, 2001; Chen e Guestrin, 2016).

O RF é um método de aprendizado de máquina do tipo *ensemble*, baseado na combinação de múltiplas árvores de decisão treinadas de forma independente, cujo princípio central consiste na redução da variância do modelo por meio da agregação de previsões (Breiman, 2001). O algoritmo utiliza duas fontes principais de

aleatoriedade: a amostragem *bootstrap* dos dados de treinamento (*bagging*) e a seleção aleatória de subconjuntos de variáveis a cada divisão dos nós das árvores. Essa estratégia reduz a correlação entre as árvores individuais, tornando o modelo final mais robusto a ruído, *outliers* e variações nos dados, características comuns em aplicações industriais.

Cada árvore de decisão no RF é construída por meio de partições recursivas do espaço de atributos, buscando minimizar uma função de impureza, como o erro quadrático médio no caso de regressão. A previsão final do modelo é obtida pela média das previsões das árvores individuais. Uma vantagem importante do *Random Forest* é sua capacidade de generalização, aliada à relativa insensibilidade ao ajuste fino de hiperparâmetros, o que facilita sua aplicação prática. Além disso, o método permite estimar a importância das variáveis, fornecendo subsídios para análise do processo e interpretação dos fatores mais relevantes para a predição (Breiman, 2001; Russell e Norvig, 2021). Contudo, embora mais interpretável do que redes neurais profundas, o RF perde transparência quando comparado a uma única árvore de decisão, e pode apresentar desempenho inferior a métodos de *boosting* em problemas altamente complexos.

O *Extreme Gradient Boosting (XGBoost)* é um algoritmo avançado baseado em *gradient boosting*, no qual árvores de decisão são treinadas sequencialmente, de modo que cada nova árvore busca corrigir os erros cometidos pelas anteriores (Chen e Guestrin, 2016). Diferentemente do *Random Forest*, o *XGBoost* atua principalmente na redução do viés, ajustando progressivamente um modelo aditivo que minimiza uma função de perda diferenciável.

Uma característica central do *XGBoost* é a incorporação explícita de termos de regularização na função objetivo, penalizando a complexidade das árvores e contribuindo para maior capacidade de generalização. Além disso, o algoritmo inclui mecanismos eficientes de paralelização, tratamento de dados ausentes e controle da profundidade das árvores, o que o torna altamente escalável e adequado para bases de dados de grande porte e alta dimensionalidade. Em aplicações industriais, o *XGBoost* tem se destacado pelo alto desempenho preditivo, frequentemente superando outros métodos baseados em árvores, embora apresente maior

sensibilidade ao ajuste de hiperparâmetros e maior custo computacional (Chen e Guestrin, 2016; Ge *et al.*, 2017).

Do ponto de vista interpretativo, o *XGBoost* é menos transparente do que o *Random Forest*, uma vez que o modelo final resulta da combinação sequencial de múltiplas árvores dependentes entre si. Ainda assim, técnicas de análise de importância de variáveis e métodos explicáveis complementares podem ser empregados para mitigar essa limitação em aplicações industriais críticas (Pensa *et al.*, 2025).

Na indústria de celulose estas técnicas foram utilizadas em diferentes partes do processo. Shan, Liu e Tang (2022) usaram rede neural e regressão logística para prever resíduos alcalinos na polpa após a lavagem e o grau Baumé (°Bé) do licor preto. Karlström, Hill e Johansson (2024) também usaram rede neural artificial para prever propriedades da polpa, como largura, comprimento da fibra e grau de drenabilidade. He e colaboradores (2022) empregaram *Random Forest*, *Gradient Boosting Regression*, *Ridge Regression* e *K-nearest neighbor* para monitorar a resistência ao dobramento, força de ruptura, suavidade e resistência à compressão na direção transversal e longitudinal em um processo de fabricação de papel.

É válido destacar que existe uma infinidade de técnicas de aprendizado de máquina. Descreveu-se nesta seção aquelas que apresentaram os melhores resultados conforme o objetivo proposto neste trabalho.

### **3.3 Revisão bibliométrica**

Foi realizada uma revisão bibliométrica a fim de identificar as tendências de pesquisa sobre o forno de cal da indústria de celulose ao longo do tempo. Os dados foram coletados a partir de uma pesquisa na base Scopus, que é largamente utilizada para revisões bibliográficas.

Para a obtenção de artigos relacionados ao forno de cal na indústria de celulose, utilizou-se a seguinte estratégia de busca, considerando palavras-chave presentes no título, resumo e palavras-chave: TITLE-ABS-KEY("lime kiln" OR "industry kiln" OR "rotary kiln") AND TITLE-ABS-KEY("pulp industry" OR "pulp mill" OR "cellulose" OR "forest industry"). A busca foi realizada em janeiro de 2025 e considerou publicações

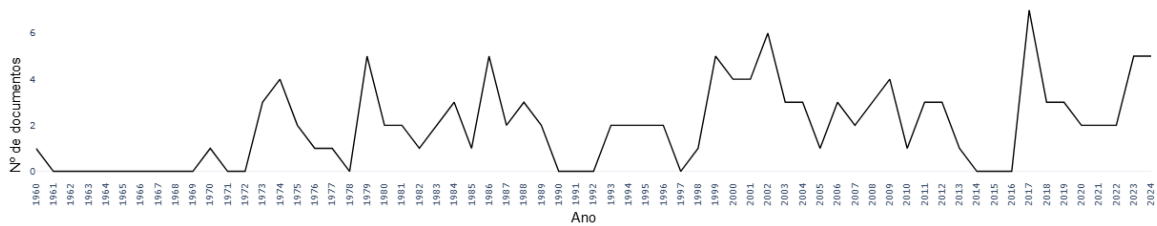
até 2024. Não foram considerados resenhas, artigos e resenhas de conferências, capítulos de livros, livros, notas, pesquisas curtas, teses, dissertações e relatórios.

Em seguida, foi realizada uma triagem manual composta pela leitura de títulos e resumos para identificar e excluir artigos que não se enquadravam nos objetivos desta revisão. Uma leitura completa foi realizada quando necessário para eliminar quaisquer dúvidas. O banco de dados foi examinado para eliminar registros de referências duplicadas, informações ausentes e artigos que não se concentravam em fornos de cal. O conjunto de dados final foi usado para avaliar padrões de publicação, tendências de autores, redes de palavras-chave e a evolução da literatura científica sobre fornos de cal na indústria de celulose. A análise bibliométrica foi realizada usando a função “biblioshiny” do pacote “bibliometrix” no Software R (Aria e Cuccurullo, 2017). Essa ferramenta foi utilizada para analisar a produção associada aos autores, a produção científica por país, as tendências ao longo do tempo e a citação do artigo. O VOSviewer, versão 1.6.20 (Van Eck e Waltman, 2010) foi usado para construir e visualizar redes de co-ocorrência de palavras-chave.

A busca original resultou em 234 artigos, dos quais 125 atenderam aos critérios de seleção e foram incluídos ( $\approx 53\%$ ) e 109 ( $\approx 47\%$ ) não atenderam aos parâmetros estabelecidos e foram excluídos.

Uma análise do número de publicações por ano (Figura 5) demonstrou que a maioria dos picos corresponde a anos após marcos ambientais significativos, como a Conferência de Estocolmo (1972), o Protocolo de Kyoto (1997), o Acordo de Paris (2015) e a COP26 (2021). Essa correlação sugere que a pesquisa no campo dos fornos de cal está intimamente ligada à sustentabilidade e às preocupações ambientais. As flutuações na produção de publicações durante as décadas de 1970 e 1980 também podem ser atribuídas ao ritmo acelerado do desenvolvimento industrial durante esse período, juntamente com a crise global do petróleo, que aumentou o interesse em eficiência energética e fontes alternativas de combustível. Publicações que abordam desempenho energético (Hsieh, 1983; Scobczynski, 1984) e fontes de combustível (Azarniouch e Philp, 1983; Philp e Azarniouch, 1984; McCubbin, 1985; Barquin e Monsarro 1986) corroboram essa ideia.

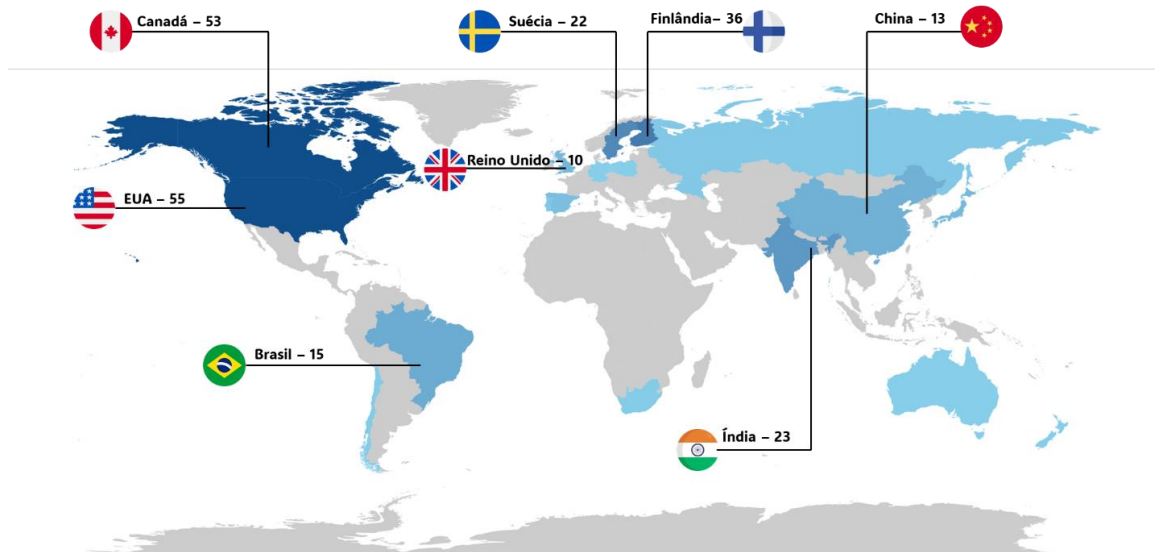
**Figura 5: Frequência de publicação por ano relacionada a estudos sobre fornos de cal na indústria de celulose.**



Fonte: Autoria própria.

O mapa na Figura 6 ilustra a contribuição global para o campo. Embora a pesquisa tenha sido realizada em 18 países, os Estados Unidos emergem como o principal contribuinte, o que pode ser um reflexo da base industrial de longa data do país e do forte financiamento acadêmico e de pesquisa. Os EUA são seguidos pelo Canadá, Finlândia, Índia, Suécia, Brasil, China e Reino Unido. Com exceção da Índia e do Reino Unido, todos os países estão entre os dez maiores produtores globais de celulose (Ibá, 2025). A inclusão da Índia é justificada por seu status como o mercado de papel que mais cresce em todo o mundo, com o consumo doméstico de papel para embalagens e papelão aumentando a uma taxa anual de 8,2% em 2023–24, de acordo com a *Indian Paper Manufacturers Association* (IPMA). O país também é um dos maiores produtores de papel do mundo.

**Figura 6: Visão geral da pesquisa mundial sobre fornos de cal na indústria de celulose.**



Fonte: Autoria própria.

Apesar de ter pouca ou nenhuma produção de celulose e, portanto, nenhum forno de cal, e se concentrar principalmente na produção de papel, o Reino Unido é um dos

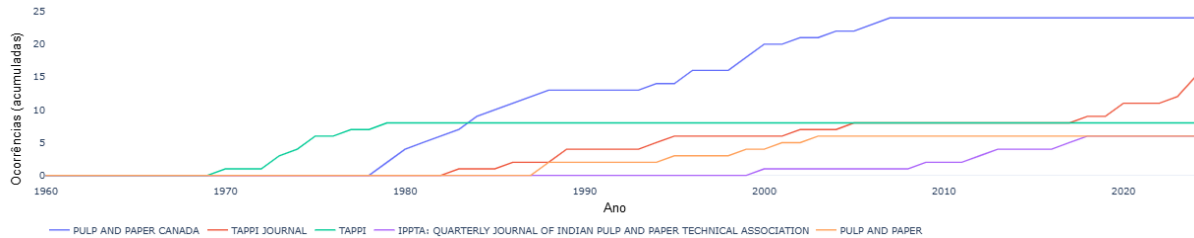
principais contribuintes para a publicação acadêmica. Isso pode ser devido à sua forte ênfase na educação e à presença de inúmeras grandes instituições, onde a produção de pesquisa é naturalmente alta. Notavelmente, apesar de ser o segundo maior produtor mundial (Ibá, 2025), o Brasil contribuiu com apenas 15 artigos em 65 anos, enquanto a China, o terceiro maior produtor, contribuiu com 13 artigos. Tal diferença entre produção de celulose e produção científica sobre fornos de cal nos dois países sugere que esses países têm oportunidades significativas de aumentar suas contribuições para publicações globais. A Finlândia é também um contribuinte notável, refletindo a indústria florestal como o principal ator no setor industrial do país (Lipiäinen e Vakkilainen, 2021).

A Finlândia e o Canadá também possuem as universidades com as afiliações mais relevantes: a Universidade de Tecnologia Lappeenranta-Lahti (LUT) na Finlândia, com treze publicações, e a Universidade de Toronto (U of T) no Canadá, com seis publicações. Embora a Espanha não esteja entre os países que mais publicam, a Universidade de Oviedo está logo atrás da LUT e da U of T nas afiliações institucionais mais proeminentes, com um total de cinco publicações. O autor com maior número de artigos publicados entre os identificados nos documentos analisados, Vakkilainen, E., também é filiado à Finlândia. Ele é seguido por quatro autores afiliados a institutos canadenses: Tran, H., Azarniouch, M. K., Murray, F.E. e Prakash, C.B. Esses resultados reforçam a importância desses dois países para a pesquisa sobre fornos de cal na indústria de celulose.

A análise da produção das fontes ao longo do tempo, apresentada na Figura 7, sugere que as publicações ocorrem, em sua maioria, em periódicos específicos da área de celulose e papel. Isso indica o foco de uma indústria em plataformas dedicadas para disseminar pesquisas e sugere que os avanços e discussões críticas para o setor geralmente se concentram em publicações do nicho. A Associação Técnica da Indústria de Celulose e Papel (TAPPI) tem um papel importante: antes de 1980, era a principal editora como organização e, nos anos seguintes, a Revista TAPPI registra o maior número de publicações, com um aumento acentuado nos últimos anos, especialmente na última década. Embora não haja registro de publicações sobre fornos de cal na *Pulp and Paper Canada* nos últimos anos, foi uma fonte importante até os anos 2000 e possui o maior número de ocorrências acumuladas. Desde 2010,

o *Indian Pulp & Paper Technical Association* (IPPTA) tem experimentado um crescimento gradual, justificado pelo crescimento na publicação de artigos pela Índia.

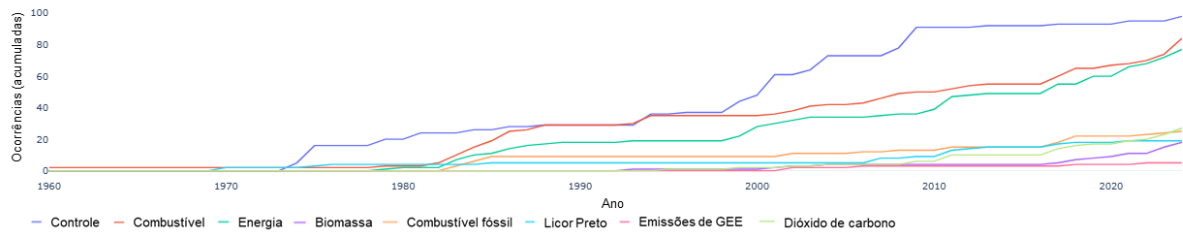
**Figura 7: Acúmulo progressivo de publicações sobre fornos de cal na indústria de celulose por fonte ao longo do tempo.**



Fonte: Autoria Própria.

Uma análise da frequência das palavras presentes nos resumos ao longo do tempo (Figura 8) sugere um foco no controle do processo e nas preocupações ambientais. Isso indica que o controle do forno de cal tem sido uma preocupação, especialmente em relação a problemas operacionais, como formação de anéis e bolas (Tran, Mao e Barham, 1993; Lindblom e Theliander, 1999; Francey e Tran, 2012; Mohanan *et al.*, 2018). Em seus estágios iniciais, o controle estava relacionado principalmente a operações manuais e empíricas. Muitas dessas publicações estão associadas ao controle de odores nas fábricas, que está relacionado às emissões de TRS (Francey e Tran, 2012; Prakash, Mathur e Murray, 1973; Thakore, 1979; Trauffer, 1995; Järvensivu, Kivivasara e Saari, 1999; Hruška *et al.*, 2020). Por exemplo, Prakash, Mathur e Murray (1973) estudaram a secagem da lama de cal e a oxidação de seu sulfeto de sódio usando um pré-secador de leito de jorro para controlar as emissões de sulfeto de hidrogênio. Bergmann e Rying (1975) relataram uma melhoria na operação do forno de cal controlando a deposição de materiais orgânicos e inorgânicos por meio de métodos de base química, como a aplicação de polímeros. Mas algumas publicações iniciais já demonstraram relevância para conceitos agora associados a sistemas de controle e automação modernos. Uronen e Aurasmaa (1979) usaram modelos estatísticos e dinâmicos para simular um forno de cal e propor estratégias de controle usando um computador digital. Scobczynski (1984) relatou um estudo de viabilidade sobre a redução do uso de energia por meio da automação que incluía fornos de cal.

**Figura 8: Acúmulo progressivo de publicações associadas a tópicos-chave de pesquisa na área de forno de cal.**



Fonte: Autoria própria.

À medida em que a pesquisa avançou, novas ferramentas e tecnologias foram implementadas. Osmond, Tessier e Savoie (1994) desenvolveram e compararam duas estratégias baseadas em controladores integrais proporcionais de *feedback* (PI), desacopladores e *loops feedforward* para apresentar a dinâmica de um forno industrial. Carter e Rozek (2000) apresentaram um analisador de carbonato residual virtual comissionado em 1998. No ano seguinte, Järvensivu, Saari e Jämsä-Jounela (2001) já usaram modelos de redes neurais, lógica difusa e equação linguística para construir um sistema inteligente de controle de forno. Em 2024, uma câmera de triangulação a laser *in-situ* foi usada para monitorar a distribuição de tamanho de nódulos de cal (Berg *et al.*, 2024). A fluidodinâmica computacional (CFD) e o aprendizado de máquina também são usados para prever variáveis e controlar fornos rotativos de cal, conforme relatado em Ryan, Bussmann e DeMartini (2022). Este trabalho também aborda preocupações com a temperatura do casco, um aspecto crítico no campo para garantir a operação confiável e eficiente do equipamento.

Os estudos mencionados anteriormente demonstram que o controle de fornos rotativos evoluiu significativamente de métodos clássicos para abordagens inteligentes baseadas em dados. A integração de aprendizado de máquina, sensores avançados e ferramentas de monitoramento em tempo real reflete uma tendência mais ampla na automação industrial em direção à manutenção preditiva, controle aprimorado de processos e eficiência operacional. No geral, a pesquisa sobre sistemas de controle geralmente se sobrepõe aos esforços para melhorar a eficiência energética, indicando uma forte correlação entre essas áreas.

Em relação aos assuntos mais tratados, “combustível” é o segundo termo mais frequente nos resumos e apresenta forte crescimento, especialmente após 2000, confirmando a tendência do setor de substituir os combustíveis fósseis. Dez

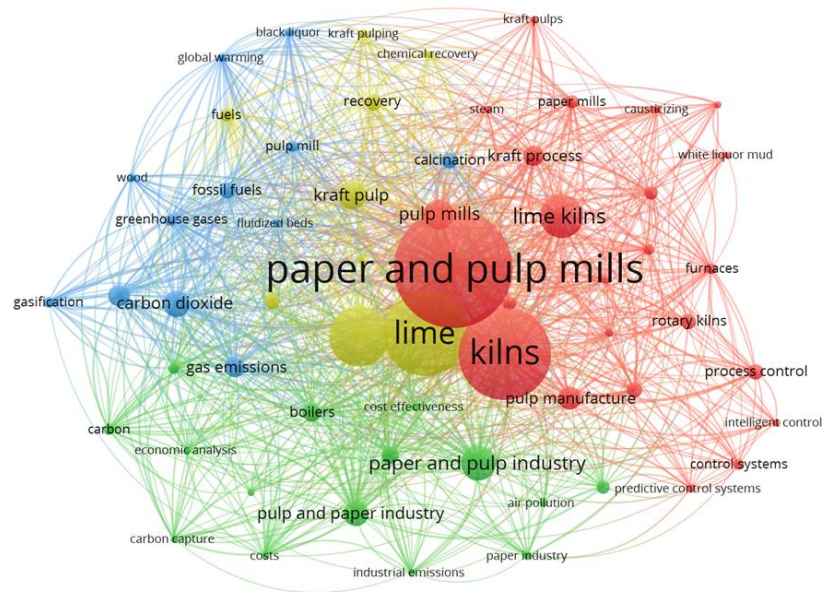
ocorrências em resumos foram registradas apenas em 2024. Como esperado, "energia" tem um padrão semelhante, com crescimento contínuo, uma vez que estão relacionadas. "Combustível fóssil" apresenta ocorrências mínimas até o início dos anos 2000, o que ocorre após o Protocolo de Kyoto (1997), o primeiro acordo internacional para reduzir as emissões de gases de efeito estufa. Alguns registros (Barquin e Monsarro, 1986; Manning, Cooper e Macfadyen, 2003) mostram que o uso de coque de petróleo foi considerado como combustível alternativo na indústria de papel devido ao aumento dos custos de energia, mesmo após o protocolo. Por outro lado, foram identificados artigos que abordaram diferentes fontes de combustível (McCubbin, 1985), incluindo madeira (Azarniouch e Philp, 1983; Philp e Azarniouch, 1984), o que sugere que as fábricas de celulose têm uma prática de longa data de utilizar uma mistura diversificada de combustíveis, incluindo subprodutos gerados a partir de seus próprios processos. O uso crescente de termos como "biomassa", "emissões de efeito estufa" e "dióxido de carbono" nos últimos anos reflete o foco crescente nas mudanças climáticas e nas preocupações ambientais.

Em 2017, três dos sete artigos publicados foram sobre combustíveis renováveis, especialmente biocombustíveis, e dois abordaram a captura de CO<sub>2</sub>. Este é também o ano com maior número de publicações, como destacado anteriormente. Isso pode ser um reflexo do Acordo de Paris, assinado em 2015 e em vigor desde 2016, que visa limitar o aquecimento global bem abaixo de 2°C acima dos níveis pré-industriais, com esforços para limitá-lo ainda mais a 1,5°C. Após a COP26, quando as nações foram instigadas a adotar metas mais ambiciosas de redução de emissões para 2030 e alinhar essas metas com o alcance de emissões líquidas zero até meados do século, muitos artigos relacionados a biocombustíveis e descarbonização foram publicados (Rey, Pio e Tarelho, 2021; Ahmadvand e Sowlati, 2023; Lefvert e Grönkvist, 2023; Martinez *et al.*, 2023; Khadivi e Sowlati, 2024).

Os *clusters* temáticos identificados na análise de palavras-chave (Figura 9) sugerem um forte foco em eficiência energética, controle de emissões e otimização de processos na indústria de papel e celulose em relação ao forno de cal. Foi definido um número mínimo de cinco ocorrências por palavra-chave, resultando em um total de 60 termos incluídos na rede, os quais foram organizados em quatro clusters. O tamanho do nó representa a frequência de cada termo no conjunto de dados,

enquanto a proximidade entre os nós reflete a força de sua co-ocorrência. A espessura das linhas de conexão indica a frequência com que os termos aparecem juntos. Os clusters, mostrados em cores diferentes, destacam áreas temáticas distintas dentro da literatura.

**Figura 9: Rede de interações entre palavras-chave indexadas e de artigos.**



Fonte: Autoria própria.

O *cluster* vermelho inclui "fornos", "controle de processo", "controle inteligente" e "sistemas de controle preditivo", sugerindo um foco no controle e otimização do processo. O cluster verde está relacionado à sustentabilidade e análise econômica e inclui "captura de carbono", "emissões industriais", "poluição do ar", "análise econômica" e "custo-benefício". O *cluster* azul indica pesquisas sobre fontes de energia, combustíveis e redução de emissões. Termos como "gaseificação", "gases de efeito estufa" e "combustíveis fósseis" estão nesta parte da rede. O *cluster* amarelo parece um cluster de conexão. Ele destaca a interseção de eficiência energética, redução de emissões e viabilidade econômica nas operações industriais.

Foram identificadas algumas lacunas com base nas publicações mais recentes. O impacto das condições operacionais na qualidade da cal ainda precisa ser mais investigado (Berg *et al.*, 2024), especialmente considerando as novas tecnologias disponíveis, incluindo inteligência artificial e aprendizagem de máquina. Eles também

podem ajudar a entender melhor o comportamento transiente do forno (Ryan, Bussmann e DeMartini, 2024), que pode ser explorado com ferramentas como CFD. Outro campo que precisa de atenção é a captura de CO<sub>2</sub>. Técnicas inovadoras, como a oxidação e a calcinação por plasma de arco elétrico, devem ser aprimoradas e testadas em escala industrial no setor de celulose. A oxidação consiste na queima de combustíveis em um ambiente rico em oxigênio, em vez de ar atmosférico, o que resulta em gases de exaustão com alta concentração de dióxido de carbono, facilitando sua captura e reduzindo as emissões de poluentes. Já a calcinação por plasma de arco elétrico utiliza um arco elétrico para gerar um plasma de alta temperatura, capaz de promover a decomposição térmica de materiais carbonáceos ou minerais de forma mais eficiente e controlada, com potencial para reduzir o consumo energético e as emissões associadas aos processos convencionais (Lefvert e Grönkvist, 2023).

## 4 VISITAS TÉCNICAS

As visitas técnicas são reconhecidas como ferramentas pedagógicas importantes no processo de ensino-aprendizagem, uma vez que possibilitam a integração entre o conhecimento teórico e a aplicação prática, favorecendo a compreensão de conceitos abstratos e complexos (Cavalcante *et al.*, 2024; Akbar *et al.*, 2025; Oliveira *et al.*, 2025; Sarema e Yan, 2026). Além disso, são oportunidades para o fortalecimento das relações entre a academia e o setor produtivo, promovendo a colaboração que pode impulsionar a inovação, o desenvolvimento tecnológico e a geração de novos conhecimentos (Carbone *et al.*, 2020).

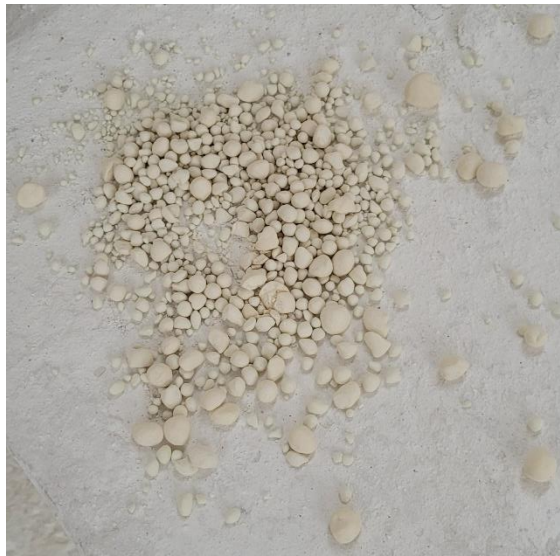
Durante o desenvolvimento deste trabalho, foram realizadas visitas técnicas a quatro fábricas de celulose, sendo três localizadas na região Sudeste (uma em Minas Gerais e duas no estado de São Paulo) e uma na região Nordeste do Brasil. Dentre elas, apenas uma apresentava integração com a produção de papel no mesmo complexo industrial; as demais dedicavam-se exclusivamente à fabricação de celulose, destinando 100% da polpa produzida ao processamento em outras unidades. Embora todos os processos industriais tenham sido observados, o foco das visitas concentrou-se nos fornos de cal e nas Estações de Tratamento de Efluentes (ETEs), áreas diretamente relacionadas às linhas de pesquisa desenvolvidas pelos alunos orientados pelo professor responsável por este trabalho.

Todas as visitas técnicas incluíram o acompanhamento de engenheiros e operadores durante a rotina operacional e mostraram-se complementares entre si. Em algumas delas, dedicou-se mais tempo à observação e discussão nos painéis de controle, enquanto em outras a explanação em campo, junto aos equipamentos, foi mais detalhada. Essas visitas foram fundamentais para aprofundar a compreensão dos processos industriais e proporcionar uma visualização prática dos conceitos teóricos. Além disso, as interações com operadores e gestores permitiram conhecer os desafios cotidianos da área e compreender as particularidades operacionais de cada empresa.

A visita a diferentes fábricas mostrou-se especialmente enriquecedora, pois possibilitou o contato tanto com equipamentos mais antigos quanto com tecnologias mais recentes, como os fornos com correntes e os sistemas de secagem externa

conhecidos como *Lime Mud Dryer* (LMD), respectivamente. Além disso, a diversidade das unidades visitadas permitiu comparar os processos produtivos e identificar quais práticas eram padronizadas e quais representavam adaptações específicas às condições e à realidade operacional de cada local. A Figura 10 foi registrada durante uma das visitas técnicas e apresenta uma amostra de cal coletada diretamente de um forno de cal, evidenciando as características visuais típicas de um processo de calcinação bem conduzido.

**Figura 10: Amostra de cal coletada em um forno de cal da indústria de celulose.**



Fonte: Acervo pessoal.

Além disso, as visitas técnicas foram fundamentais para o fortalecimento dos vínculos entre a indústria e a academia, promovendo uma integração mais efetiva e a identificação de oportunidades de colaboração futura que possam beneficiar ambas as partes. Essas visitas também desempenharam papel importante na viabilização da disponibilização de dados industriais para fins acadêmicos, os quais foram utilizados neste trabalho e continuarão a subsidiar estudos posteriores.

Dessa forma, as visitas técnicas contribuíram diretamente para o alcance dos objetivos deste trabalho, oferecendo subsídios empíricos e referências práticas fundamentais para a análise e discussão dos resultados apresentados nos capítulos seguintes. A escolha de parâmetros do modelo, por exemplo, foi influenciada pelas conversas com operadores e pela comparação entre a operação de diferentes fábricas. Foram priorizadas variáveis comuns a todas as instalações.

## 5 METODOLOGIA

Neste capítulo são apresentados os procedimentos metodológicos adotados para o desenvolvimento do modelo de predição de carbonato residual, incluindo a descrição das etapas de tratamento de dados, seleção de variáveis, técnicas de modelagem empregadas e pós-processamento, além das análises ambiental e econômica posteriores.

### 5.1 Estudo de caso

Para o desenvolvimento do trabalho foram utilizados dados de processo coletados em uma fábrica de celulose localizada no Nordeste brasileiro. O sistema de calcinação desta indústria é equipado com a tecnologia LMD, que promove uma melhor secagem da lama antes de entrar no forno de cal, reduzindo a área de secagem dentro do equipamento e aumentando a eficiência energética. Segundo especialistas e operadores da área, o tempo de residência da lama no forno de cal estudado é de aproximadamente 2,5 horas.

As informações referentes ao processo foram extraídas do sistema PIMS (*Plant Information Management System*) da empresa. A base de dados a ser utilizada contém informações de operação contínua do forno em um tempo de 694 dias entre março de 2023 e janeiro de 2025. Para a construção dos modelos preditivos, têm-se 19 variáveis de entrada à disposição, listadas a seguir, coletadas pontualmente em intervalos de 30 minutos:

1. Lama alimentada no filtro (m<sup>3</sup>/h)
2. Produção do Forno (t/dia)
3. Consistência da Lama (%)
4. Oxigênio residual da queima (%)
5. Fração de Hidrogênio (%)
6. Fração de Gás Natural (%)
7. Fração de Metanol (%)
8. Energia específica total (MJ/t)
9. Temperatura do Leito (°C)
10. Temperatura do Ar Secundário (°C)
11. Temperatura do Ar Primário (°C)

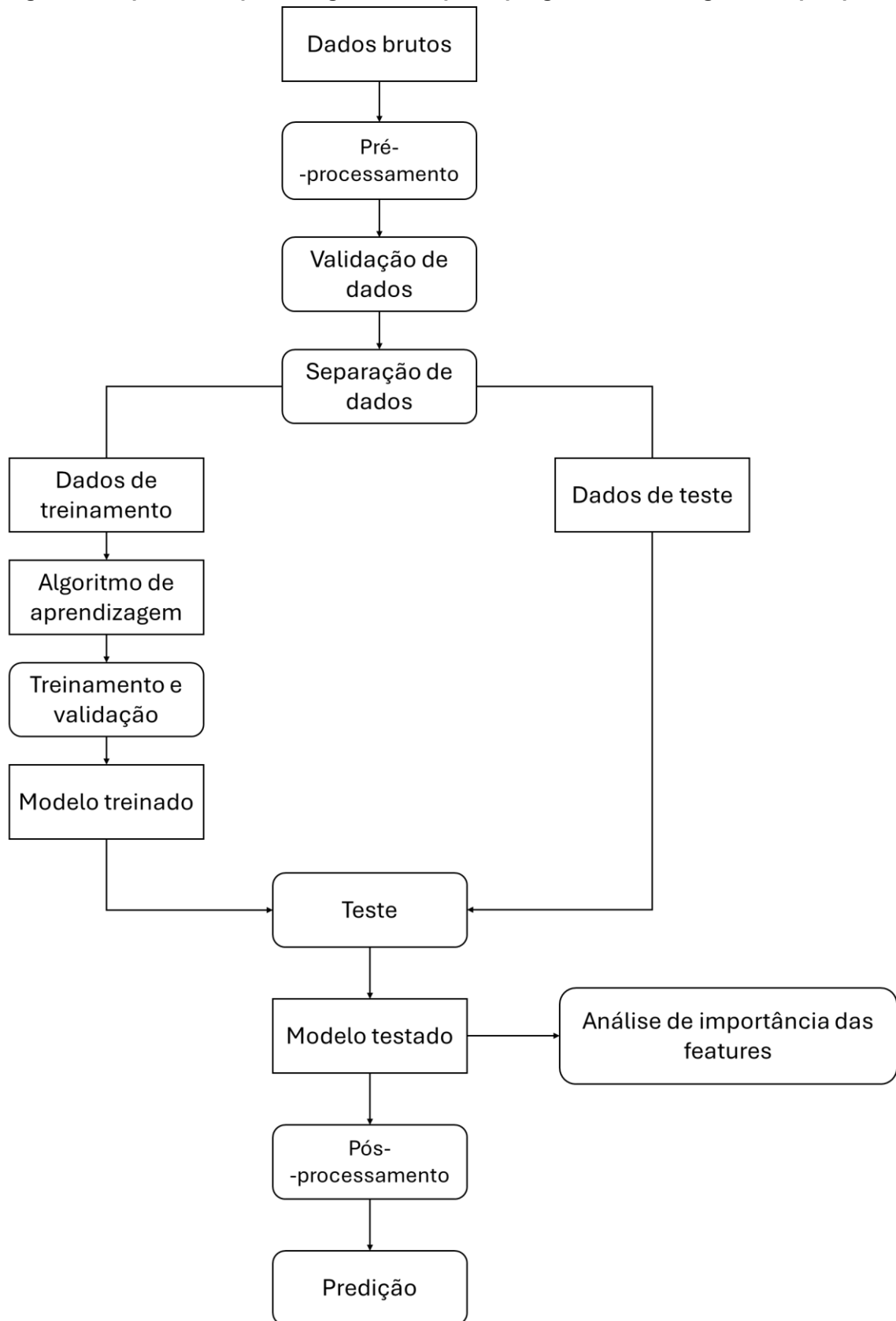
12. Temperatura dos gases de saída (°C)
13. Velocidade Ventilador (%)
14. Concentração de Monóxido de Carbono (CO) – Chaminé (ppm)
15. CaO Útil (%)
16. Teor de Carbonato de Licor Verde (g/L)
17. Dosagem de Cal de reposição(t/h)
18. Eficiência Caustificação (%)
19. *Ratio* de Caustificação (kg/m<sup>3</sup>)

A variável de saída é o teor de carbonato residual (%) após a calcinação, que é diretamente relacionado à eficiência da reação e à qualidade da lama (Tran, 2008). Atualmente, sua quantificação é feita em laboratório a cada seis horas em média, totalizando quatro medições diárias. A base de dados contém, 1.659 amostras para o carbonato residual e 1.117 dados faltantes. Os parâmetros podem ser divididos em dois grupos, o primeiro relacionado à operação do forno em si (1 a 14) e o segundo relacionado à operação de toda a planta de caustificação (15 a 19).

A Figura 11 ilustra o *pipeline* de aprendizado de máquina adotado neste trabalho, sintetizando de forma esquemática as principais etapas envolvidas no desenvolvimento do modelo, desde o pré-processamento dos dados até a obtenção e avaliação do modelo preditivo. Essa representação permite uma visão integrada do fluxo metodológico empregado, facilitando a compreensão da sequência de procedimentos e das decisões adotadas ao longo do estudo.

Todas as análises foram feitas com o uso da linguagem de programação Python e *Jupyter Notebook* como ambiente de desenvolvimento integrado. A escolha do Python se justifica por se tratar de uma linguagem gratuita, flexível, simples e possuir um vasto sistema de bibliotecas. Além disso, destaca-se por sua ampla aplicação em problemas de engenharia industrial e análise de dados (Fatoba *et al.*, 2025; Wei, 2025). Já o *Jupyter Notebook* permite a execução sequencial de códigos, bem como o registro integrado de resultados e visualizações. Essa abordagem favorece a reprodutibilidade dos experimentos e a organização do fluxo metodológico adotado no estudo (Kluyver *et al.*, 2016).

Figura 11: Pipeline de aprendizagem de máquina que guia a metodologia desta pesquisa.



Fonte: Autoria própria.

## 5.2 Pré-processamento

O pré-processamento de dados bem-sucedido depende da integração de técnicas de análise de dados e conhecimento do processo. No entanto, trabalhar com medições reais de processo traz um desafio adicional, visto que elas têm muito ruído, heterogeneidade e abrangem várias escalas. Portanto, esta etapa foi primordial para que os objetivos do trabalho fossem alcançados.

Para entender melhor os dados, foi realizada uma análise exploratória. Inicialmente foram avaliadas algumas estatísticas descritivas com o objetivo de compreender o comportamento das variáveis envolvidas no processo. Foram calculadas média, mediana, desvio padrão, valores máximos e mínimos e quartis (25%, 50%, 75%). Além disso foram plotados histogramas, séries temporais e gráficos de dispersão e de correlação cruzada. Isso permitiu a identificação de padrões, tendências, *outliers* e correlações entre as variáveis, fornecendo informações importantes para as etapas subsequentes.

Inicialmente, uma limpeza com base em conhecimentos básicos e de processos foi conduzida para eliminar registros errôneos (por exemplo, fluxos negativos e teor de carbonato residual zero) e os períodos de inatividade, partida e parada do forno. Em seguida, os *outliers* restantes foram detectados usando o algoritmo *IsoForest*, um detector de anomalias preciso e eficiente baseado em modelos parciais (Liu, Ting e Zhou, 2008). Os dados foram suavizados com o filtro Savitzky-Golay, que remove o ruído por uma aproximação polinomial de mínimos quadrados locais (Savitzky e Golay, 1964), adaptado para suavizar dados com base apenas nos dados presentes e passados. Esses métodos foram aplicados a cada variável individualmente, e qualquer amostra contendo pelo menos um valor discrepante foi descartada, resultando em um conjunto de dados de trabalho.

É importante salientar que a preservação de um grau mínimo de variabilidade nos dados é fundamental para assegurar a robustez e a aplicabilidade do modelo preditivo em cenários reais. Em ambientes industriais, é comum que os instrumentos de medição estejam sujeitos a ruídos, desvios e incertezas. Portanto, o modelo deve ser capaz de lidar com essas imperfeições sem comprometer seu desempenho. Um pré-processamento excessivo, que elimine demasiadamente as variações naturais dos

dados, pode resultar em um modelo superajustado a condições ideais e pouco representativas da realidade operacional. Isso comprometeria sua capacidade de generalização, tornando-o menos eficaz diante de situações ruidosas ou imprecisas, que são frequentes no contexto de processos industriais.

Duas abordagens de modelagem baseada em dados foram desenvolvidas neste trabalho. Na primeira, utilizada como base, não foi considerado o tempo de residência. Nesse caso, tanto as variáveis *on-line* quanto a análise de carbonato residual foram consideradas no mesmo momento da amostragem. Com o objetivo de avaliar o efeito do tempo de residência, Järvensivu, Saari e Jämsä-Jounela (2001) utilizaram médias móveis com janelas de 150 a 180 minutos aplicadas a variáveis selecionadas, permitindo representar a dinâmica lenta do forno e atenuar as variações de alta frequência. Seguindo esta ideia e com base em especialistas no forno utilizado como estudo de caso neste trabalho, na segunda abordagem foi assumido um tempo médio de residência de 2,5 horas.

Na operação industrial analisada, o forno é controlado para manter esse valor-alvo, ajustando continuamente a velocidade de rotação em resposta a variações na taxa de alimentação da lama de cal. Além disso, o conteúdo de carbonato residual usado para o desenvolvimento do modelo é obtido a partir de análises laboratoriais realizadas em intervalos aproximados de 6 horas, representando efeitos acumulativos do processo em vez de condições instantâneas do forno. Para garantir consistência entre entradas e saídas do modelo, as variáveis de entrada foram definidas como valores medianos das 2,5 horas anteriores a cada medição laboratorial. Dada essa agregação temporal e a estratégia de controle da planta, não se espera que flutuações de curto prazo no tempo de residência afetem significativamente o carbonato residual medido.

### 5.3 Desenvolvimento do modelo

Para encontrar um bom modelo preditivo baseado em dados para o teor de carbonato residual no forno de cal, foram consideradas três técnicas diferentes de aprendizado de máquina: *RandomForest* (RF); *Rede Neural Artificial* (RNA), especificamente *Multilayer Perceptron* (MLP); e *Extreme Gradient Boosting* (XGBoost). Foi utilizada a biblioteca *Scikit-learn* do *Python*, especificamente as funções “*RandomForestRegressor*”, “*sklearn.neural\_network.MLPRegressor*” e

“*XGBRegressor*”, respectivamente. Todas essas técnicas têm sido utilizadas em aplicações industriais, conforme relatado na literatura e no capítulo 3.2.3 deste trabalho.

Considerando a relevância física e operacional dos valores extremos de carbonato residual, foi adotada uma estratégia para garantir sua presença nos conjuntos de treinamento. Os valores mínimo e máximo globais foram identificados e separados antes da divisão dos dados para impedir que fossem alocados exclusivamente nos subconjuntos de validação. Caso isso ocorresse, estes registros extremos eram automaticamente realocados para o conjunto de treinamento do respectivo *fold*. Essa abordagem reduz o risco de extrapolações indevidas e contribui para a estabilidade do modelo em toda a faixa de valores de carbonato residual.

Excluídos os extremos, os dados restantes foram divididos em conjuntos de treino/validação (80%) e teste (20%). Para garantir representatividade estatística, a divisão foi realizada de forma estratificada, a partir da discretização da variável resposta em quantis utilizando cinco classes. Isso garante que a distribuição dos valores de carbonato residual fosse semelhante tanto no conjunto de treinamento/validação quanto no conjunto de teste, reduzindo vieses associados a não representatividade. O conjunto de teste foi mantido completamente isolado até a etapa final de avaliação dos modelos, para garantir uma representação satisfatória da capacidade de generalização de cada um. A ordem temporal foi desconsiderada porque a frequência dos dados era menor do que o necessário para refletir a influência temporal.

Os modelos foram implementados por meio de *pipelines* e as variáveis de entrada foram normalizadas utilizando o método Min-Max *Scaling* para limitar os dados no intervalo [0,1]. A adoção deste método para todos os modelos avaliados teve como objetivo garantir condições comparáveis de entrada, ou seja, foi uma estratégia de padronização metodológica para comparação entre modelos de naturezas distintas.

A variável resposta foi transformada por meio da função logarítmica “*log1p*” com inversão pela função *expm1*. Essa transformação foi incorporada ao processo de treinamento utilizando “*TransformedTargetRegressor*”. O objetivo foi reduzir assimetrias, estabilizar a variância e melhorar o desempenho dos modelos.

A otimização dos hiperparâmetros foi realizada por meio do método de *Grid Search*, que avalia todas as combinações especificadas no espaço de busca, selecionando aquelas que apresentam melhor desempenho de acordo com uma métrica pré-estabelecida. Para cada modelo, foi definido um conjunto de combinações de hiperparâmetros relevantes, descritas na Tabela 1. Neste trabalho, a métrica principal utilizada durante a otimização foi a raiz do erro quadrático médio (*Root Mean Squared Error* - RMSE), calculado a partir do erro quadrático médio retornado pelo procedimento de validação cruzada, conforme a Equação 4, (Newhart, Hering e Cath, 2022). Nesta notação,  $y_i$  é o valor observado,  $\hat{y}_i$  é o valor predito e  $n$  é o número de observações. Esta é uma medida amplamente empregada em problemas de regressão que penaliza erros de maior magnitude.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4)$$

**Tabela 1: Valores testados para os hiperparâmetros de cada modelo.**

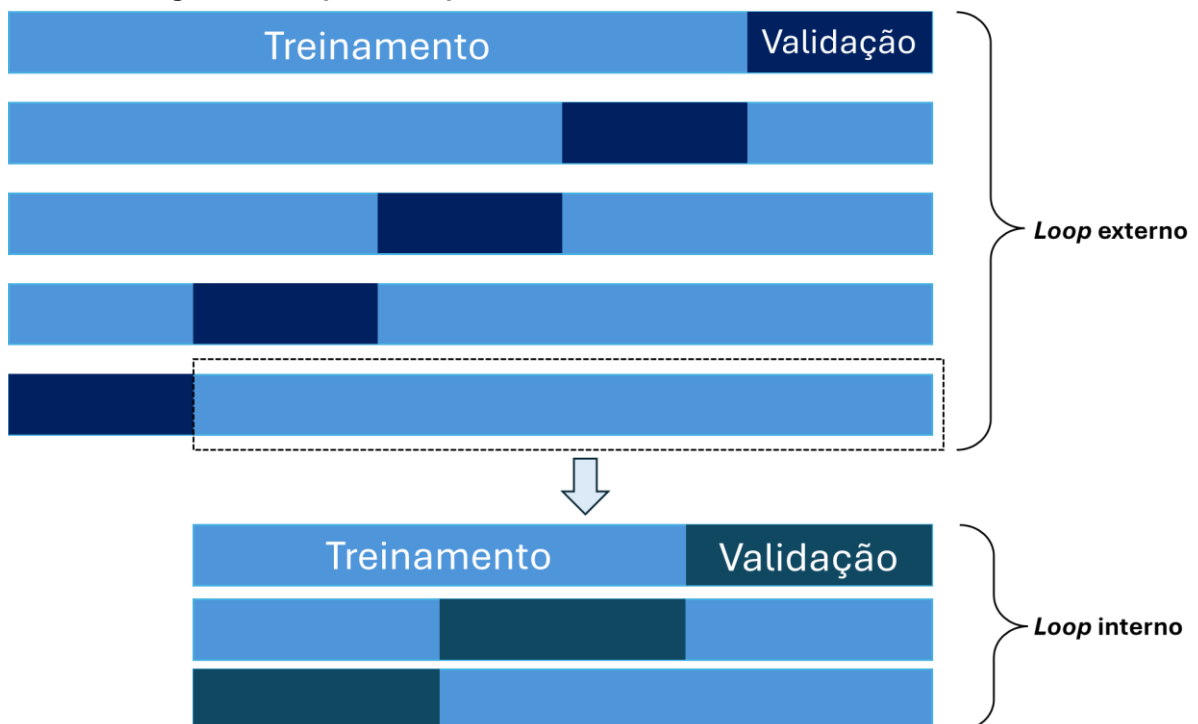
<b>Modelo</b>	<b>Hiperparâmetros</b>	<b>Valores</b>
<i>Random Forest</i>	Número de estimadores (árvores)	5 a 250, de 5 em 5
	Profundidade máxima	0 a 100
	Número mínimo de amostras em um nó	2,5
Rede Neural	Função de ativação	ReLu, Tangente hiperbólica
	Número de camadas ocultas	1, 2
	Número de neurônios na camada oculta	1 a 9
	Taxa de aprendizagem	0,001 e 0,01
	Alfa (intensidade de regularização L2)	0,0001 e 0,001
<i>XGBoost</i>	Número de estimadores (árvores)	100 a 1500, de 100 em 100
	Profundidade máxima	3 a 10
	Taxa de aprendizagem	0,01; 0,03; 0,1 e 0,3
	Fração de amostras	0,8; 0,9 e 1
	Fração de variáveis	0,8; 0,9 e 1

Fonte: Autoria própria.

Para obter estimativas de desempenho mais confiáveis e evitar vazamento de dados e um viés otimista, foi adotada a validação cruzada aninhada (*nested cross-validation*). Neste procedimento, representado na Figura 12, o conjunto de treinamento foi dividido em dois níveis, em que o *Grid Search* é executado exclusivamente no *loop* interno, enquanto o *loop* externo é responsável por avaliar a capacidade de generalização do modelo. Mais especificamente, no *loop* externo o conjunto de treinamento/validação é dividido em cinco *fold*s independentes, sendo cada *fold* utilizado alternadamente como conjunto de validação. No *loop* interno, cada subconjunto de treinamento do *fold* externo passa por uma nova validação cruzada, definida em três *fold*s para este trabalho, na qual o *Grid Search* identifica os parâmetros ótimos.

Esse procedimento garante que a escolha de hiperparâmetros seja sempre realizada com base apenas nos dados de treinamento, enquanto o desempenho reportado nos *fold*s externos reflete uma estimativa mais realista da capacidade de generalização dos modelos. O desempenho de cada *fold* externo foi avaliado por meio do RMSE, sendo reportados a média e o desvio padrão, fornecendo uma estimativa da capacidade de generalização dos modelos.

Figura 12: Esquema de partição de dados da validação cruzada aninhada.



Fonte: Autoria própria.

Após a conclusão da validação cruzada aninhada, cada modelo foi retreinado utilizando todo o conjunto de treinamento/validação com os melhores hiperparâmetros identificados. Em seguida, o desempenho final foi avaliado no conjunto de teste, que não participou de nenhuma etapa de treinamento ou otimização.

Além do RMSE utilizado no *pipeline*, também foram avaliadas as métricas de erro absoluto médio (*Mean Absolute Error* - MAE) e coeficiente de correlação ( $r$ ), conforme as Equações 5 e 6, respectivamente (Newhart, Hering e Cath, 2022). Também foram realizadas análises gráficas por meio de diagramas Predito versus Real, histogramas de resíduos e gráficos de resíduos em função de valores preditos, o que permitiu uma avaliação visual da qualidade do ajuste e da presença de vieses.

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (5)$$

$$r = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 (\hat{y}_i - \bar{\hat{y}})^2}} \quad (6)$$

Nas equações,  $\bar{y}$  é a média dos valores observados e  $\bar{\hat{y}}$  é a média dos valores preditos.

#### 5.4 Pós-processamento

Esta etapa avalia o impacto da suavização de sinal (feita na etapa de pré-processamento de dados com o uso do filtro de *Savitzky-Golay*) no desempenho do modelo. Para isso, os valores previstos do modelo final selecionado foram comparados com os valores-alvo originais ( $y_i^*$ ), sem suavização, conforme a Equação 7, na qual  $res_i$  representa o resíduo associado à observação  $i$ , indicando o erro de previsão do modelo em relação ao valor real.

Foram definidas faixas de previsão ( $B_k$ ), com largura fixa de 0,5%, iniciando em 0%. Para cada faixa foi calculado o viés médio ( $\mu_k$ ), conforme a Equação 8. Tal valor representa o erro sistemático do modelo para previsões pertencentes à faixa e  $N_k$  é o número de amostras em cada faixa. Os valores previstos do modelo foram então ajustados conforme a Equação 9, dando origem a uma previsão corrigida ( $\tilde{y}_i$ ).

$$res_i = y_i^* - \hat{y}_i \quad (7)$$

$$\mu_k = \frac{1}{N_k} \sum_{i \in B_k} res_i \quad (8)$$

$$\tilde{y}_i = \hat{y}_i + \mu_k \quad (9)$$

Também foi calculado o desvio padrão de cada faixa ( $s_k$ ), que quantifica a dispersão dos resíduos em torno do viés médio, sendo utilizado como medida de incerteza local da previsão, conforme a Equação 10. A partir da previsão corrigida, foram construídos intervalos empíricos de incerteza correspondentes a um e dois desvios-padrão, definidos, respectivamente, pelas Equações 11 e 12.

$$s_k = \sqrt{\frac{1}{N_k - 1} \sum_{i \in B_k} (res_i - \mu_k)^2} \quad (10)$$

$$\tilde{y}_i \pm s_k \quad (11)$$

$$\tilde{y}_i \pm 2s_k \quad (12)$$

Esses intervalos foram utilizados para avaliar a capacidade do modelo em reproduzir os valores reais de carbonato dentro de limites aceitáveis de erro. Por fim, foi calculada a proporção de observações reais contidas dentro dos intervalos de  $\pm 1s$  e  $\pm 2s$  em relação à previsão corrigida. Essas métricas, representadas por  $P_{1s}$  e  $P_{2s}$ , são definidas pelas Equações 13 e 14, respectivamente:

$$P_{1s} = \frac{1}{n} \sum_{i=1}^n I(|y_i^* - \hat{y}_i| \leq s_k) \quad (13)$$

$$P_{2s} = \frac{1}{n} \sum_{i=1}^n I(|y_i^* - \hat{y}_i| \leq 2s_k) \quad (14)$$

em que  $I(\cdot)$  é a função indicadora, que assume valor unitário quando a condição é satisfeita e zero caso contrário.

Essa abordagem permitiu avaliar o efeito da suavização do sinal sobre a qualidade das previsões considerando o comportamento local dos resíduos ao longo das diferentes faixas de valores previstos de carbonato.

## 5.5 Análise de importância das features

Esta etapa busca determinar o efeito de cada variável de entrada do modelo selecionado como o mais satisfatório sobre o teor de carbonato residual. Foram aplicados três métodos para avaliar diferentes aspectos das contribuições dos

preditores. O *F-Score* fornece a influência relativa de cada variável na estrutura do modelo (Scikit-learn Developers, 2025a). A métrica de Importância Nativa, própria do *XGBoost*, destaca a contribuição da variável para reduzir o erro do modelo (Scikit-learn Developers, 2025b). A permutação avalia o quanto o modelo depende de cada variável para previsões precisas (Scikit-learn Developers, 2025c). As importâncias obtidas por cada método foram normalizadas no intervalo de 0 a 1 e posteriormente combinadas por meio da média aritmética, com o objetivo de reduzir vieses inerentes a cada técnica individual e fornecer uma avaliação mais confiável da relevância das variáveis. A adoção de múltiplos métodos de análise de importância é recomendada na literatura para aplicações em processos industriais complexos, nos quais diferentes variáveis podem influenciar o sistema de maneira não linear e interdependente (Molnar, 2022; Wang *et al.*, 2024). Os resultados foram então associados às condições de operação do forno.

## 5.6 Análise ambiental e econômica

Esta etapa visa quantificar o potencial de redução de emissões de CO<sub>2</sub> associado à eliminação de condições de supercalcinação no forno em questão. O limite operacional alvo foi definido neste estudo como 3% em fração mássica, valor compatível com as práticas industriais da empresa do estudo de caso deste trabalho. Cada medição de carbonato residual foi assumida como representativa de um intervalo médio de operação de 6 horas, conforme a frequência típica de amostragem do processo. A taxa de produção ( $m_{cao,i}$  t/dia) foi convertida em massa efetiva de óxido de cálcio produzida por amostra, segundo a Equação 15:

$$m_{cao,i} = \frac{P_i}{24} \cdot \Delta t \quad (15)$$

em que  $P_i$  é a taxa de produção do forno no instante  $i$  (t/dia) e  $\Delta t$  é o intervalo de tempo representado por cada medição (6h).

A análise considera exclusivamente os instantes operacionais em que o carbonato residual medido encontra-se abaixo do valor alvo, representando situações de consumo energético e emissões reduzidas. A magnitude da sobrequeima em cada instante (*overburn<sub>i</sub>*) foi definida como a diferença entre o valor alvo de carbonato residual ( $r^*$ ) e o valor medido ( $r_i$ ), conforme a Equação 16. Valores positivos indicam

condições de sobrequeima, enquanto valores nulos representam operação dentro ou acima do alvo, não contribuindo para o cálculo das emissões reduzidas.

$$overburn_i = \max(r^* - r_i, 0) \quad (16)$$

As emissões de CO<sub>2</sub> reduzidas associadas à sobrequeima foram divididas em duas parcelas. A primeira corresponde às emissões de processo, resultantes da decomposição térmica excedente do carbonato de cálcio em relação à demanda do processo produtivo. Elas foram estimadas por meio de um fator de emissão específico, expresso em toneladas de CO<sub>2</sub> por tonelada de CaO e por unidade de sobrequeima, conforme a Equação 17, que é baseado na estequiometria.

$$CO_{2,proc,i} = K_{proc} \cdot overburn_i \quad (17)$$

em que  $K_{proc}$  é o coeficiente estequiométrico associado à liberação de CO<sub>2</sub> durante a supercalcinação, considerado como 0,785 a partir da Equação 3.

A segunda parte corresponde às emissões de combustão, associadas ao consumo excessivo de combustível necessário para sustentar temperaturas acima do necessário. Primeiramente, a energia térmica evitável ( $E_i$ ) associada à eliminação da sobrequeima foi estimada a partir de um coeficiente energético específico ( $K_E$ ), que é a energia da reação expressa em Gigajoule por tonelada de CaO e por unidade de sobrequeima. Considerando que a entalpia da reação descrita na Equação 3 é de 178kJ, o valor de  $K_E$  foi assumido como 3,17GJ/tCaO, e não foram consideradas as perdas térmicas do forno. Assim, a energia térmica foi calculada conforme a Equação 18. As emissões de CO<sub>2</sub> provenientes da combustão ( $CO_{2,comb,i}$ ) foram então calculadas conforme a Equação 19. Foi considerado o fator de emissão (EF) do gás natural, visto que é o único combustível fóssil utilizado pela empresa, que é de 56,1kgCO<sub>2</sub>/GJ (IPCC, 2006).

$$E_i = K_E \cdot overburn_i \quad (18)$$

$$CO_{2,comb,i} = \frac{E_i \cdot EF}{1000} \quad (19)$$

A emissão total de CO<sub>2</sub> reduzida (CO<sub>2,tot,i</sub>) por tonelada de CaO em cada instante é dada pela soma das parcelas de processo e combustão, conforme a Equação 20:

$$CO_{2,tot,i} = CO_{2,proc,i} + CO_{2,comb,i} \quad (20)$$

Os valores absolutos de emissões reduzidas foram obtidos multiplicando-se as emissões específicas pela massa de CaO representada por cada amostra de acordo com a Equação 21.

$$CO_{2,tot} = \Sigma[(CO_{2,proc,i} + CO_{2,comb,i}) \cdot m_{CaO,i}] \quad (21)$$

Na avaliação econômica foram quantificados os benefícios monetários potenciais associados à eliminação das condições de sobrequeima no processo de produção de cal. Foram consideradas a economia direta de combustível, decorrente da diminuição do consumo energético do forno, e a valoração econômica das emissões de CO<sub>2</sub> reduzidas, a partir de preços praticados no mercado voluntário de carbono.

A energia evitada calculada anteriormente foi convertida em valor monetário segundo a Equação 22 utilizando um preço médio de gás natural (C<sub>f</sub>). Tal valor foi definido como 18,956US\$/GJ (20US\$/MMBtu) de acordo com relatório da Confederação Nacional da Indústria (2025). O valor total de economia de combustível no período analisado foi obtido pela soma das economias individuais associadas a cada instante de operação em condição de sobrequeima.

$$Economia_{comb,tot} = \Sigma(E_i \cdot C_f) \quad (22)$$

A valoração econômica das emissões de CO<sub>2</sub> reduzidas foi calculada conforme a Equação 23. Para a definição do preço unitário do carbono (C<sub>CO2</sub> CUS\$/tCO<sub>2</sub>) considerou-se os preços reportados pelo *Ecosystem Marketplace* (2025), sendo que para os cálculos foi utilizado o preço médio de créditos de carbono para projetos da categoria de Processos Químicos e Indústria Manufatureira nos anos de 2023 e 2024 de US\$3,88.

$$Valor_{CO2} = \Sigma CO_{2,i} \cdot C_{CO2} \quad (23)$$

O benefício econômico total potencial associado à eliminação da supercalcinação foi calculado pela soma da economia direta de combustível e do valor econômico das emissões de CO<sub>2</sub>, conforme a Equação 24.

$$Benefício_{tot} = Economia_{comb,tot} + Valor_{CO_2,tot} \quad (24)$$

É importante destacar que a metodologia adotada assume um cenário ideal de controle, no qual as condições de supercalcinação seriam eliminadas nos instantes em que ocorrem, sem afetar negativamente a qualidade do produto final. Dessa forma, os resultados representam um potencial de mitigação, condicionado à implementação de estratégias de controle adequadas. Para um cálculo real seriam necessárias a implementação do modelo e a quantificação de redução de emissões de acordo com padrões reconhecidos para a creditação de carbono. Apesar dessas limitações, a metodologia permite quantificar de forma consistente os benefícios ambientais e econômicos associados à otimização do processo, fornecendo subsídios para decisões técnicas e gerenciais.

## **6 RESULTADOS E DISCUSSÃO**

Neste capítulo são apresentados e discutidos os resultados obtidos a partir da aplicação da metodologia proposta.

### **6.1 Estudo de caso**

A seleção criteriosa das variáveis que seriam consideradas como entrada foi essencial para garantir que o modelo fosse construído a partir de informações relevantes, estáveis e representativas do comportamento físico do sistema. Para o presente trabalho o grupo de variáveis relacionadas à etapa de caustificação foi desconsiderado, pois o tempo de sua influência na operação do forno de cal é incerto e a natureza de sua correlação não é clara. Além disso, a consistência da lama também foi desconsiderada por ser uma medida laboratorial, não contribuindo para a construção de um sensor virtual com variáveis online.

Para reduzir a dimensionalidade do problema, variáveis relacionadas ao forno foram selecionadas com base no conhecimento do processo. O objetivo era capturar os principais aspectos físicos, químicos e operacionais do processo do forno, sem tornar o modelo excessivamente específico a uma única planta. Dessa forma, ele apresenta potencial de adaptação para implementação em diferentes ambientes industriais, considerando que nem todas as instalações monitoram os mesmos parâmetros. Também se esperava facilitar a análise de importância das variáveis e sua relação com aspectos do processo.

Ao final, nove variáveis (Tabela 2) foram consideradas como entradas para o modelo. Juntas, abrangem as entradas (combustível, ar e alimentação), as condições internas do processo (temperaturas e vazão) e os indicadores de desempenho (composição do gás de combustão e eficiência energética), tornando-as adequadas para prever os níveis de carbonato residual em tempo real. A variável resposta é o teor de carbonato (%).

Tabela 2: Variáveis selecionadas como entrada para o modelo.

Variável	Unidade
Energia específica	MJ/t
Temperatura do leite	°C
Temperatura do ar primário	°C
Temperatura do ar secundário	°C
Temperatura dos gases de saída	°C
Velocidade do ventilador	%
Oxigênio residual	%
Vazão de lama (para o filtro)	m <sup>3</sup> /h
Concentração de CO (na chaminé)	Ppm

Fonte: Autoria própria.

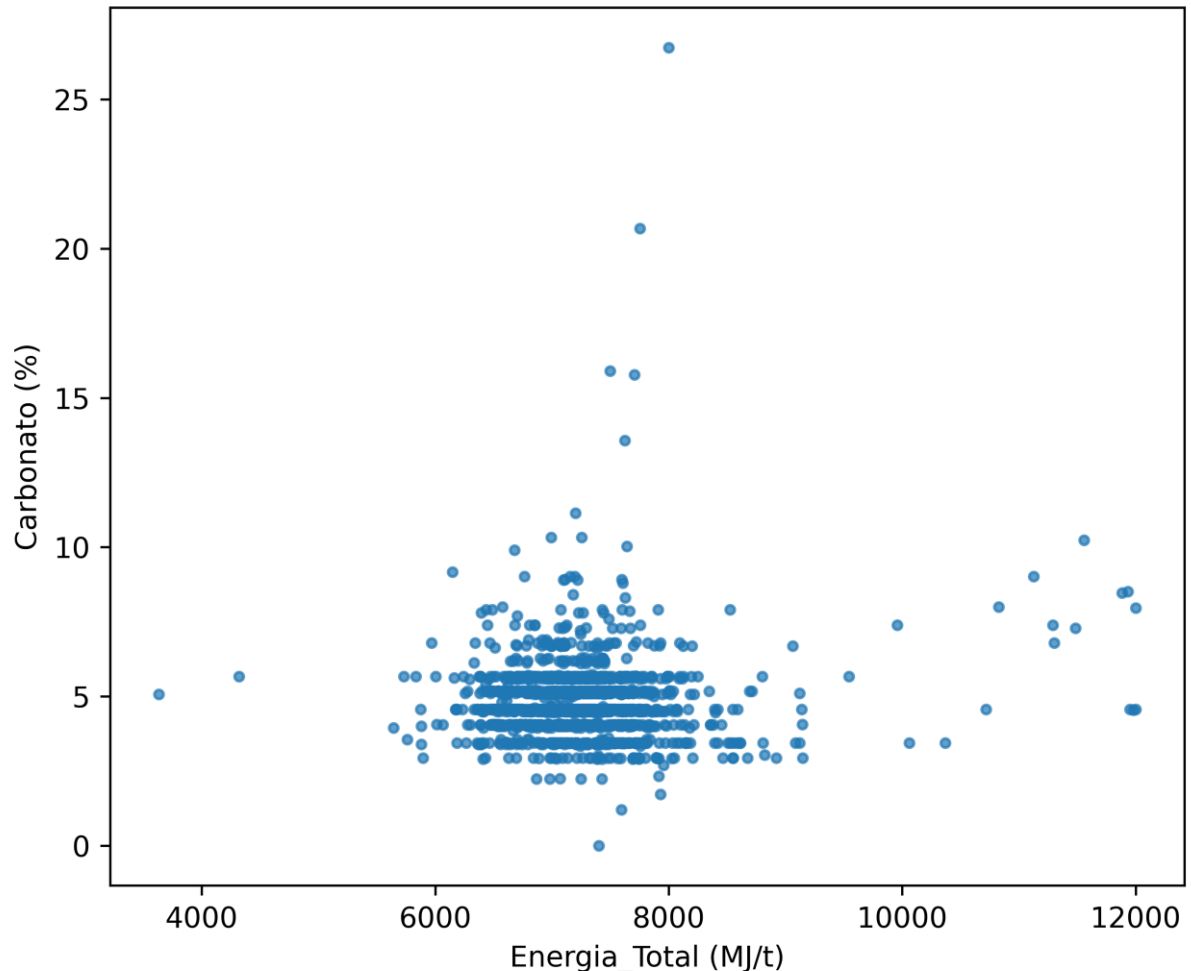
## 6.2 Pré-processamento

A etapa inicial de análise descritiva demonstrou que, conforme esperado, os dados brutos provenientes da operação industrial apresentavam valores fisicamente inconsistentes, além da presença de ruídos e outliers. A Figura 13 ilustra essa situação ao mostrar o gráfico de dispersão entre a energia específica total consumida e o teor de carbonato, no qual se observa uma concentração predominante dos dados em uma faixa operacional típica tanto para o carbonato, em torno de 3% a 6%, quanto para a energia, aproximadamente entre 6 000 e 8 000 MJ/t. Esses valores contrastam com pontos isolados que indicam comportamentos atípicos. Esses valores discrepantes, encontrados em todos os parâmetros, corroboraram a necessidade de aplicação de técnicas de pré-processamento para garantir a qualidade e a representatividade dos dados utilizados na modelagem.

Além disso, ainda observando a Figura 13, não se observa uma relação clara entre a energia e o teor de carbonato. Não há uma tendência monotônica evidente de aumento ou diminuição do carbonato com a energia, mas sim uma relação que parece ser multivariada e não-linear, apesar de saber-se que a quantidade de energia no forno é diretamente relacionada à ocorrência da reação. Além disso, a presença de *outliers* significativos contribui para o aumento da dispersão e dificulta a identificação de uma relação direta entre as variáveis. Tais comportamentos foram observados em todos os gráficos de dispersão plotados. Isto sugere que vários fatores operacionais

podem exercer influência relevante sobre o teor de carbonato, reforçando novamente a necessidade de pré-processamento dos dados e da consideração conjunta de múltiplas variáveis na etapa de modelagem.

**Figura 13: Dispersão entre Energia Específica e Carbonato Residual com dados brutos.**



Fonte: Autoria própria.

Após a análise inicial dos dados, foi realizada uma remoção de outliers conhecendo parâmetros físicos e operacionais do forno. Foram desconsiderados os pontos em que a produção do forno era menor que 500t/dia, por corresponderem a períodos de regime transiente do forno, provavelmente antes de entrar em parada ou em retomada de produção. Também foram desconsideradas medições de carbonato residual igual a zero, sendo possivelmente um erro de digitação ou algum momento em que não houve análise laboratorial. Dessa forma, a base de dados passou a conter 1.598 registros de carbonato residual, uma redução de aproximadamente 3,7% dos dados iniciais desta variável.

Após a remoção “manual” de *outliers*, foi aplicado o algoritmo *IsoForest* com um parâmetro de *contamination* igual a 0,026, resultando em um *silhouette* score de 0,7720, o que indica uma boa separação entre os dados considerados normais e os *outliers* identificados. Em relação à etapa de suavização dos sinais, a melhor configuração do filtro de *Savitzky–Golay* foi obtida com uma janela de cinco pontos e um polinômio de terceira ordem. A Figura 14 apresenta um recorte dos dados de carbonato residual antes da aplicação do *IsoForest* e já suavizados, evidenciando o efeito do filtro. Observa-se que essa combinação promoveu uma redução do ruído sem comprometer as informações relacionadas à tendência e à variabilidade do processo de operação do forno (Figura 14a). Por outro lado, ao se utilizar uma janela de onze pontos e um polinômio de terceira ordem (Figura 14b), a suavização tornou-se excessiva, resultando na perda de variações importantes e de características relevantes do processo.

**Figura 14: Dados suavizados (recorte) - (a)Janela de 5 pontos. (b)Janela de 11 pontos.**



Fonte: Autoria própria.

O pré-processamento adequado foi essencial para que o objetivo geral de se obter um bom modelo para a predição de carbonato fosse atingido, pois garantiu a qualidade das informações utilizadas na modelagem.

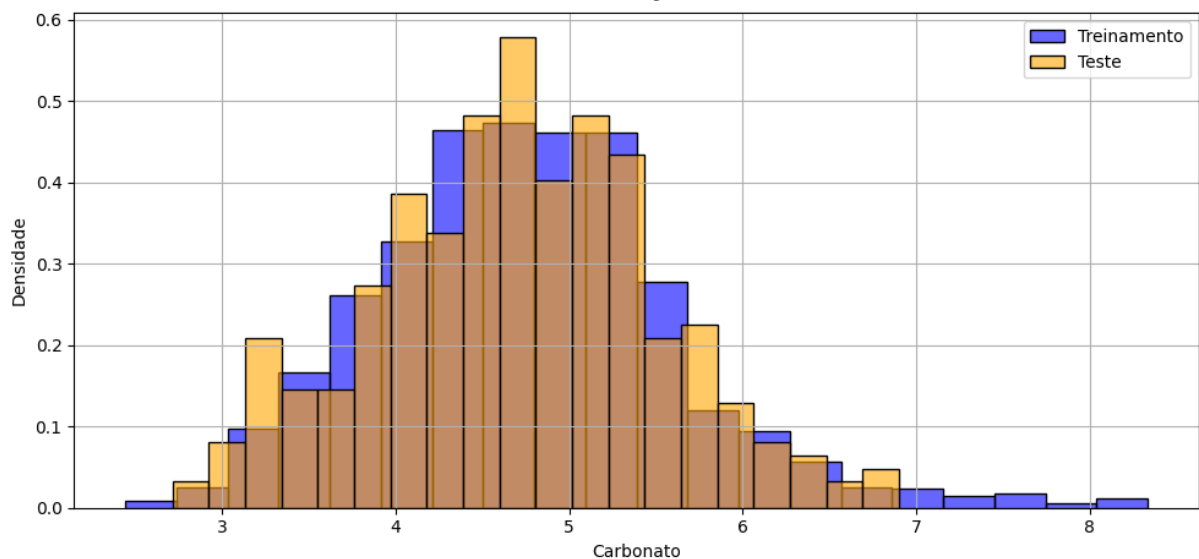
O conjunto de dados de trabalho resultante continha 1.483 amostras para a abordagem pontual e 1.524 amostras para a abordagem do tempo de residência. Isso

representa uma redução de 10,6% e 8,1% em relação ao número inicial de 1659 amostras residuais de carbonato nos dados brutos fornecidos pela fábrica, respectivamente.

### 6.3 Desenvolvimento do modelo

A separação estratificada dos dados resultou em conjuntos de treinamento/validação e teste com distribuições semelhantes do teor de carbonato residual, conforme evidenciado na Figura 15. Observa-se que a faixa de variação e a concentração dos valores foram equilibradas entre os conjuntos, garantindo os extremos no treinamento, indicando que o procedimento adotado foi eficaz em garantir representatividade estatística e em reduzir possíveis vieses decorrentes de distribuições assimétricas na avaliação do desempenho dos modelos.

**Figura 15: Distribuição estratificada de Carbonato Residual nos conjuntos de treinamento/validação e teste.**



Fonte: Autoria própria.

Os melhores resultados dos modelos, avaliados por meio dos indicadores  $r$ , MAE e RMSE, para as abordagens pontual e considerando o tempo de residência, são apresentados nas Tabelas 2 e 3, respectivamente. Maiores coeficientes de correlação e menores valores de MAE e RMSE foram obtidos considerando o tempo de residência de 2,5 horas, em comparação com a abordagem pontual. Isso destaca a importância de levar em conta o tempo de residência ao prever o conteúdo residual de carbonato, conforme relatado em outros trabalhos (Berg *et al.*, 2024). Essa relevância pode estar associada ao fato de que essa abordagem permite incorporar

efeitos acumulativos e atrasos físicos que não são capturados quando o tempo de residência é desconsiderado.

**Tabela 3: Resultados do desempenho dos modelos no subconjunto de teste na abordagem pontual.**

Modelo	Hiperparâmetros	r	MAE	RMSE
MLP	Função de ativação: tanh	0,29±0.03	0,69±0.01	0,97±0.01
	Número de camadas ocultas: 2]			
	Número de neurônios na camada oculta: (8,8)			
	Taxa de aprendizagem: 0,01			
	Alfa: 0,0001			
<i>Random Forest</i>	Número de estimadores (árvores): 130	0,63±0.01	0,53±0.01	0,75±0.01
	Profundidade máxima: 95			
	Número mínimo de amostras em um nó: 2			
<i>XGBoost</i>	Número de estimadores (árvores): 1200	0,73±0.01	0,46±0.01	0,61±0.01
	Profundidade máxima: 8			
	Taxa de aprendizagem: 0,03			
	Fração de amostras:0,9			
	Fração de variáveis:0,8			

Fonte: Autoria própria.

**Tabela 4: Resultados do desempenho dos modelos no subconjunto de teste na abordagem considerando o tempo de residência.**

Modelo	Hiperparâmetros	r	MAE	RMSE
MLP	Função de ativação: ReLu	0,45±0,06	0,56±0.02	0,72±0.02
	Número de camadas ocultas: 2			
	Número de neurônios na camada oculta: (9,4)			
	Taxa de aprendizagem: 0,01			
	Alfa: 0,0001			
<i>Random Forest</i>	Número de estimadores (árvores): 160	0,80±0.01	0,36±0.01	0,48±0.01
	Profundidade máxima: 30			
	Número mínimo de amostras em um nó: 2			
<i>XGBoost</i>	Número de estimadores (árvores): 1400	0,83±0.01	0,33±0.01	0,45±0.01
	Profundidade máxima: 8			
	Taxa de aprendizagem: 0,03			
	Fração de amostras:0,8			
	Fração de variáveis:0,8			

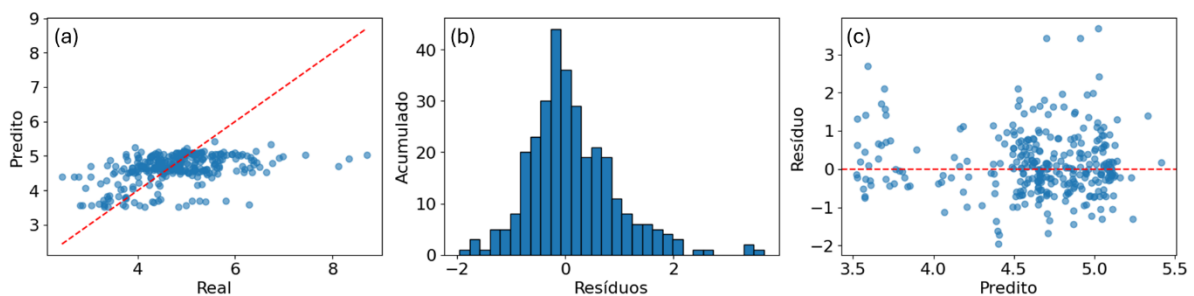
Fonte: Autoria própria.

Os modelos baseados em árvores, ou seja, *Random Forest* e *XGBoost*, apresentaram desempenho notavelmente superior ao das redes neurais. Esse resultado pode ser atribuído à maior robustez desses algoritmos frente à presença de ruídos e *outliers*,

características presentes nos dados industriais analisados, visto que certa variabilidade foi mantida. Além disso, modelos baseados em árvores tendem a lidar melhor com conjuntos de dados tabulares de tamanho moderado e com variáveis correlacionadas, demandando menor esforço de pré-processamento quando comparados às redes neurais artificiais (Breiman, 2001; Chen e Guestrin, 2016). Resultados semelhantes são reportados na literatura, que aponta que, em problemas de regressão envolvendo dados tabulares industriais, métodos baseados em *ensemble* de árvores frequentemente superam modelos neurais em termos de desempenho preditivo e estabilidade (Shavitt e Segal, 2018; Grinsztajn, Oyallon e Varoquaux, 2022).

Tais resultados também podem ser vistos na análise de resíduos e nos gráficos de comparação entre valores preditos e valores reais. Os resultados do modelo de rede neural MLP, mostrados na Figura 16, mostraram que o modelo não foi capaz de capturar muito bem a tendência dos dados (Figura 16a). Foi possível perceber uma tendência à subestimação para valores mais elevados, tipicamente acima de 5%, e à superestimação para valores mais baixos, abaixo de 5%, sugerindo viés sistemático. Apesar de apresentarem uma distribuição aproximadamente simétrica centrada em zero (Figura 16b), a dispersão observada (Figura 16c) sugere que há espaço para melhorias.

**Figura 16: Resultados do modelo MLP. (a) Valores reais vs preditos. (b) Histograma de resíduos. (c) Distribuição de resíduos em função de valores preditos.**

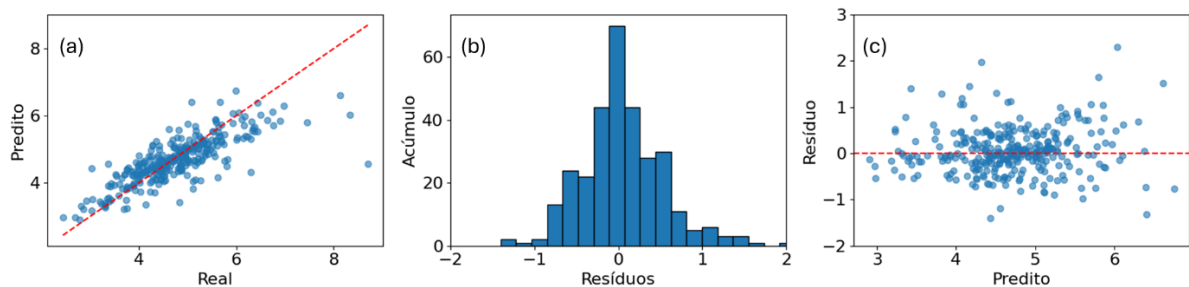


Fonte: Autoria própria.

O modelo *Random Forest* apresentou resultados notavelmente melhores que o modelo anterior, conforme mostrado na Figura 17, e com indicação de um desempenho preditivo satisfatório. Observa-se, no gráfico de valores reais versus preditos (Figura 17a), uma forte relação linear entre as variáveis, com a maioria dos pontos distribuída próxima à linha de referência, o que sugere adequada capacidade

do modelo em reproduzir os valores observados. Há uma leve tendência de superestimação de valores extremos pequenos e subestimação de valores extremos mais elevados, o que pode estar associado a um menor número de dados nestas regiões, que fogem da operação típica do forno. A distribuição dos resíduos (Figura 17b) apresenta comportamento aproximadamente normal e centrado em torno de zero, indicando ausência de viés sistemático nas predições. Além disso, o gráfico de resíduos em função dos valores preditos (Figura 17c) não evidencia padrões definidos nem heterocedasticidade pronunciada, o que reforça a suposição de que os erros estão distribuídos de forma aleatória. Esses resultados sugerem que o modelo é consistente, bem ajustado aos dados e adequado para fins de predição no contexto deste estudo.

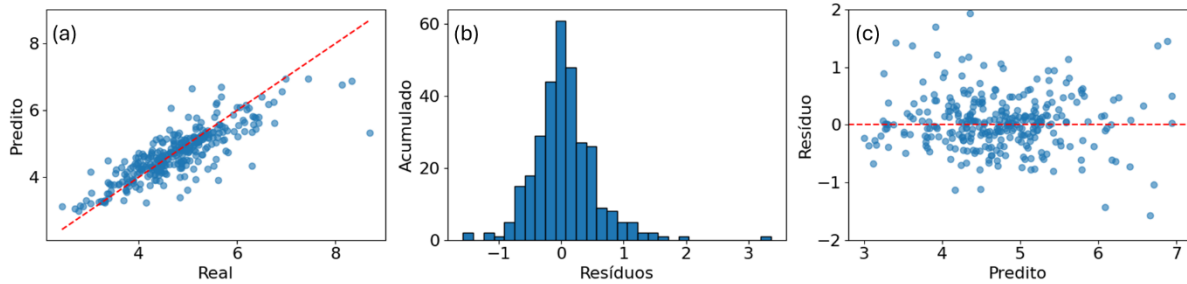
**Figura 17: Resultados do modelo *Random Forest*. (a) Valores reais vs preditos. (b) Histograma de resíduos. (c) Distribuição de resíduos em função de valores preditos.**



Fonte: Autoria própria.

As mesmas observações do RF foram feitas para o modelo *XGBoost*, cujos resultados são representados na Figura 18. No entanto, o *XGBoost* apresentou uma relação entre os valores preditos e reais ainda maior (Figura 18a). Além disso, embora seu conjunto de resíduos apresente maior amplitude, foi observada uma concentração significativamente maior dos valores em torno de zero (Figura 18b), o que indica um melhor ajuste global do modelo. A presença de poucos resíduos extremos não compromete a qualidade preditiva geral, uma vez que a maioria das observações apresenta erros reduzidos. A dispersão dos resíduos ao longo da faixa de carbonato residual (Figura 18c) também sugere aleatoriedade. Dessa forma, o *XGBoost* foi escolhido como melhor algoritmo preditivo para este estudo, o que se justifica pela maior precisão média e pela melhor representação dos dados observados. Outros trabalhos relataram o uso dessa técnica de aprendizado de máquina como ferramenta de previsão na indústria de papel e celulose (Rodrigues *et al.*, 2022; Liu *et al.*, 2025).

**Figura 18: Resultados do modelo *XGBoost*. (a) Valores reais vs preditos. (b) Histograma de resíduos. (c) Distribuição de resíduos em função de valores preditos.**

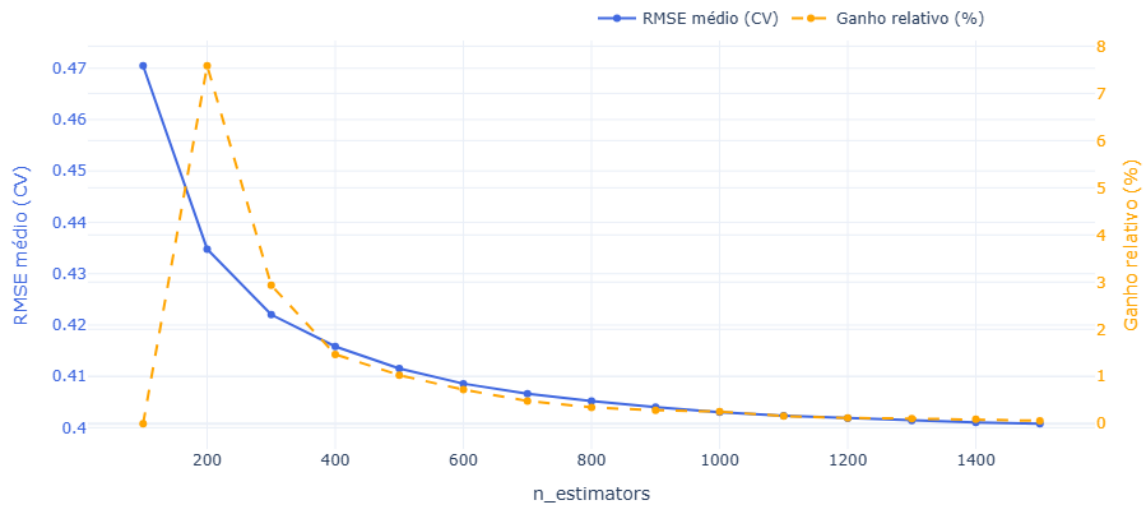


Fonte: Autoria própria.

No entanto, seguindo o princípio de parcimônia e buscando reduzir a complexidade do modelo e também garantir que não haja sobreajuste, foi realizada uma análise do desempenho do modelo em função do número de árvores. A Figura 19 apresenta a curva de validação do modelo *XGBoost* em função do número de árvores ( $n\_estimators$ ), considerando o RMSE obtido por validação cruzada, bem como o ganho relativo associado ao acréscimo de árvores ao modelo. Para essa análise, todos os demais hiperparâmetros foram mantidos fixos nos valores ótimos determinados previamente, de modo a isolar o efeito do número de estimadores no desempenho do modelo.

Observou-se que o aumento do número de árvores resultou inicialmente em uma redução do RMSE médio, indicando melhora progressiva da capacidade preditiva do modelo. Entretanto, à medida que o número de árvores aumenta, o ganho relativo entre iterações sucessivas tende a diminuir, evidenciando um comportamento de saturação do desempenho. Esse resultado sugere que, a partir de determinado ponto, o acréscimo de árvores passa a oferecer ganhos limitados.

**Figura 19: Validação do modelo XGBoost em função do número de árvores.**



Fonte: Autoria própria.

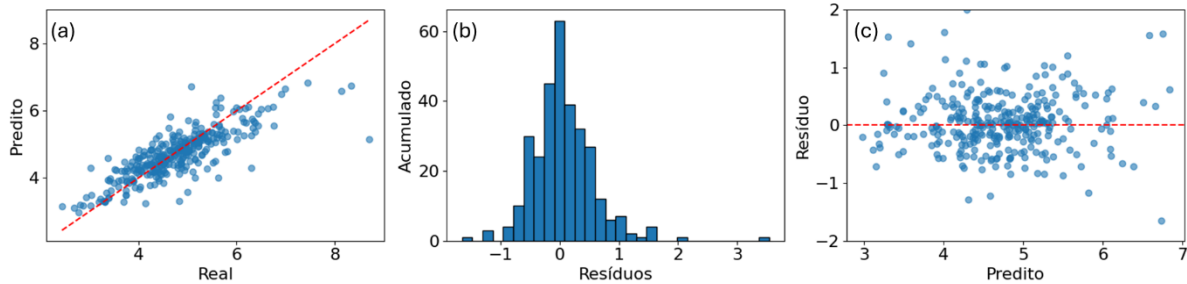
Para manter um equilíbrio entre desempenho preditivo e complexidade, um modelo com 500 árvores foi treinado com os mesmos procedimentos anteriores. Neste ponto, o ganho relativo de RMSE passou a ser menor que 1% com o aumento do número de estimadores.

Sob essa consideração, o *XGBoost* continuou gerando um resultado satisfatório, com um coeficiente de correlação consideravelmente alto, igual a  $0,83 \pm 0,01$ , sem diferença significativa para o modelo anterior. O desvio padrão relativamente baixo sugere estabilidade durante o treinamento do modelo, conforme desejável. O MAE foi igual a  $0,34 \pm 0,01\%$ , um erro aceitável para o teor de carbonato residual na prática industrial.

A Figura 20 ilustra os resultados do *XGBoost* no conjunto de dados de teste. A Figura 20a mostra a alta associação entre as previsões do modelo (eixo y) e os valores da fábrica (eixo x) para o teor de carbonato residual. As maiores diferenças apareceram no extremo superior nos valores reais da fábrica (acima de 6), que provavelmente vêm do número limitado de amostras. Como eles estão longe das condições usuais de operação do forno, tais desvios eram esperados. As Figuras 20a e 20c mostram, respectivamente, o histograma dos resíduos e a dispersão deles. É possível perceber que os resíduos apresentaram uma distribuição aproximadamente normal e um valor médio próximo de zero e são aleatoriamente distribuídos, conforme desejado. Além disso, aproximadamente 53% dos resíduos estão dentro do intervalo  $[-0,25, +0,25]$  e quase 78% dentro de  $[-0,5, +0,5]$ . Esses valores residuais relativamente pequenos

sugerem que o modelo poderia complementar as análises laboratoriais de rotina. Assim, definiu-se como melhor modelo o XGBoost com 500 árvores, utilizado nas etapas posteriores de avaliação.

**Figura 20: Resultados do modelo preditivo final. (a) Valores reais vs preditos. (b) Histograma de resíduos. (c) Distribuição de resíduos em função de valores preditos.**



Fonte: Autoria própria.

Assim, o modelo permite estimar o teor de carbonato residual antes da disponibilidade dos resultados laboratoriais, antecipando essa informação em um intervalo que, no estudo de caso analisado, variou de seis horas para trinta minutos, de acordo com o intervalo de amostragem das variáveis de entrada. Desta forma, ele pode ser utilizado como um sensor virtual para a detecção de anomalias. A detecção antecipada de desvios operacionais pode auxiliar nas atividades de diagnóstico do processo, estabelecendo uma relação mais direta entre variações no carbonato residual e suas possíveis causas. Como consequência, intervenções antecipadas na operação do forno de cal, quando necessárias, tendem a beneficiar o desempenho produtivo, os aspectos ambientais e a segurança operacional. Adicionalmente, a aplicação do modelo pode contribuir para novas percepções sobre a dinâmica de operação dos fornos de cal.

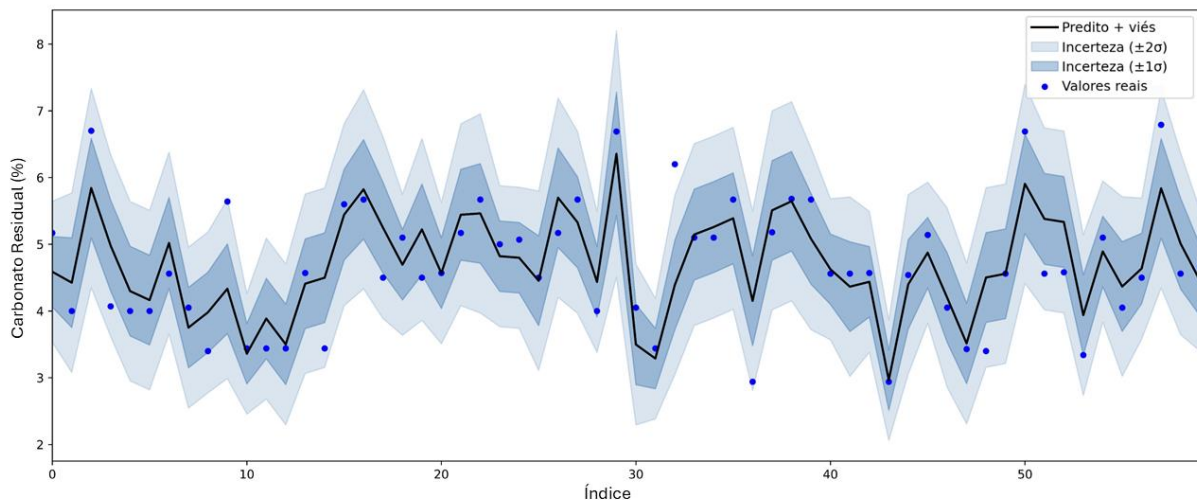
Uma outra aplicação seria voltada para o controle de processo. A integração do modelo ao sistema de controle de combustível do forno pode favorecer a redução do consumo energético, o aumento da estabilidade operacional, a melhoria da qualidade da cal produzida e a manutenção do teor de carbonato residual dentro de uma faixa ótima. Sob a perspectiva ambiental, essa integração pode contribuir para a redução das emissões de GEE provenientes da queima de combustível, especialmente  $\text{CO}_2$ , além de reduzir a liberação de produtos associados à combustão incompleta, como monóxido de carbono. No contexto do ciclo do cálcio, a aplicação do modelo pode assegurar uma qualidade de cal mais consistente para a etapa subsequente de

caustificação, com menor variabilidade, reduzindo a necessidade de purgas excessivas e de reprocessamento de material.

#### 6.4 Pós-processamento

A Figura 21 apresenta os valores previstos ajustados (linha preta) pelo modelo *XGBoost* selecionado anteriormente. Os valores originais do teor de carbonato residual, coletados na fábrica, também são mostrados (pontos azuis). Além disso, os intervalos correspondentes a um (sombra escura) e dois (sombra clara) desvios padrão também são plotados. Esses intervalos apresentam faixas empíricas de incerteza local, baseadas na dispersão observada dos resíduos. Essas bandas são úteis para quantificar a incerteza na previsão do modelo. Cerca de 73% dos dados da fábrica (pontos azuis) estão no primeiro intervalo e 94% no segundo. A ausência de padrões estruturados nos resíduos sugere que o modelo não apresenta sinais evidentes de sobreajuste. Esse resultado também sugere que o filtro de suavização Savitzky-Golay não comprometeu a precisão do modelo. Assim, o teor de carbonato residual no forno de cal foi descrito satisfatoriamente pelo modelo *XGBoost* usando um conjunto de nove variáveis de processo medidas online. A Figura 21 mostra apenas um intervalo específico de dados de teste para fins de visualização, mas o padrão registrado foi observado em todo o conjunto de teste.

**Figura 21: Ajuste do modelo para o carbonato residual, mostrando os valores observados, a série predita com correção de viés e as faixas de incerteza associadas às estimativas.**

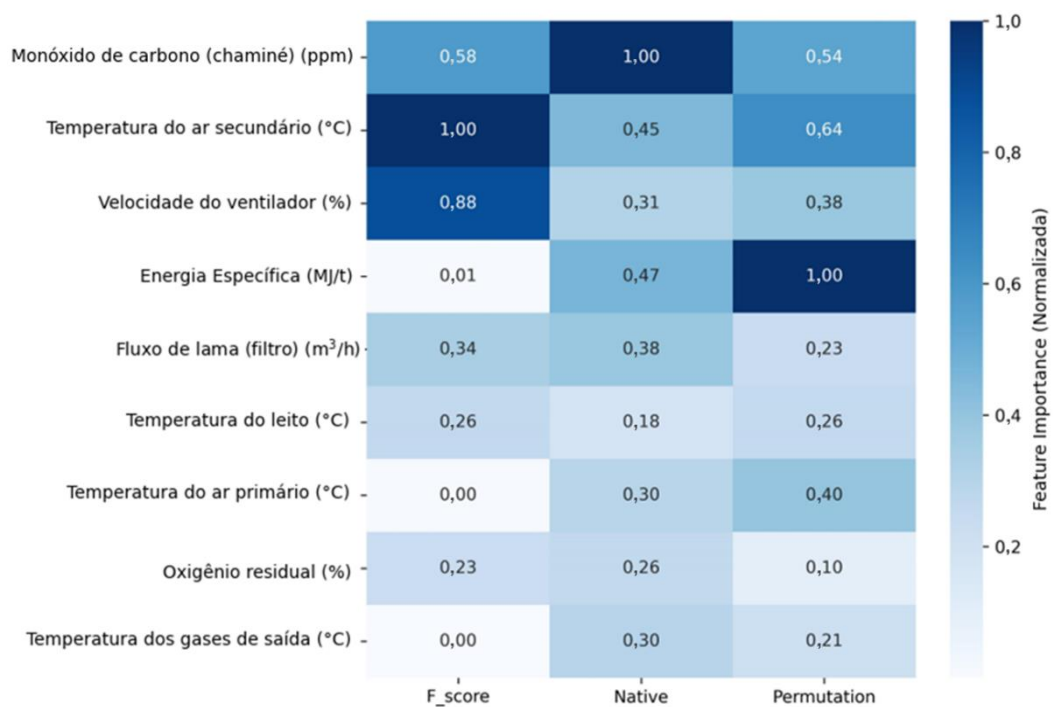


Fonte: Autoria própria.

## 6.5 Análise de importância das *features*

A análise da importância das *features* revelou que, neste caso, a concentração de CO no gás de combustão da chaminé (em ppm), a temperatura do ar secundário (em °C) e a velocidade do ventilador (em %) são as variáveis mais influentes para prever o teor de carbonato residual (em %). Essas variáveis mostram alta relevância em todas as três técnicas (Figura 22). Quanto mais próximo da unidade (quanto mais escura a cor), maior a importância da variável.

Figura 22: Análise de importância das *features*.



Fonte: Autoria própria.

A concentração de CO pode ser um indicador da eficiência da combustão. Níveis mais altos de monóxido de carbono sinalizam oxidação incompleta do combustível e deficiências de temperatura que favorecem a retenção de carbonato de cálcio (Tran *et al.*, 2007; Fernández, Turrado e Abanades, 2019; Cui *et al.*, 2023). A relevância da temperatura do ar secundário (em °C) deve-se ao seu papel governante tanto na estabilidade e conclusão do processo de combustão quanto na eficiência da transferência de calor para o leito (Teguh *et al.*, 2024). Ambos afetam diretamente a disponibilidade de energia para a decomposição do  $\text{CaCO}_3$ . A velocidade do ventilador (em %) é um fator-chave dos padrões de fluxo de ar e circulação de gás,

afetando a estabilidade da chama, a distribuição de calor e o tempo de residência do gás dentro da câmara do forno, todos os quais afetam o ambiente térmico necessário para uma calcinação eficiente (Lee, Park e Lee, 2017; Ryan, Busmann e DeMartini, 2022). Também controla a concentração de  $\text{CO}_2$  na fase gasosa.

Essas variáveis também podem ser diretamente relacionadas à cinética de calcinação, uma vez que são fortemente influenciadas pelas temperaturas e pressão parcial de  $\text{CO}_2$  (controlada pela velocidade do ventilador). A constante de taxa de reação é uma função da temperatura. Embora temperaturas mais altas sejam necessárias para iniciar a calcinação em um ambiente com alto teor de  $\text{CO}_2$ , a reação ocorre muito mais rapidamente do que em uma atmosfera com baixo teor de  $\text{CO}_2$  (Wu *et al.*, 2024). E quando a temperatura cai abaixo de  $800^\circ\text{C}$ , ocorre a reação reversa e ocorre a recarbonatação, de modo que o  $\text{CaO}$  externo reage com o  $\text{CO}_2$ , regenerando o  $\text{CaCO}_3$  (Tran, 2008).

Além disso, a energia específica (em MJ/t) parece ter um efeito forte, mas pouco claro, nas previsões do carbonato, pois exibe alto poder preditivo (alta importância na Permutação), apesar de ser usada apenas ocasionalmente e não influenciar tanto na estrutura do modelo (baixo F-Score). A taxa de alimentação de lama de cal (em  $\text{m}^3/\text{h}$ ) tem um impacto limitado nas previsões de carbonato, mas reflete o efeito fundamental da alimentação do forno, pois um grau de enchimento mais alto aumenta a demanda de energia e ar e, se não forem devidamente balanceados, levam à calcinação incompleta. Esses resultados são consistentes com a ideia de que tanto a dinâmica da entrada de material quanto os parâmetros relacionados à combustão devem ser cuidadosamente balanceados para alcançar o desempenho operacional estável ideal do forno e o teor de carbonato residual. As demais variáveis apresentaram menor poder preditivo, no entanto, permanecem importantes para a interpretabilidade, pois abrangem também os aspectos físicos, químicos e operacionais do forno.

## **6.6 Análise ambiental e econômica**

Durante o período analisado, 3,5% das medições apresentaram carbonato residual inferior a 3%, caracterizando condições de sobrequeima. Considerando apenas esses instantes, a redução total das emissões de  $\text{CO}_2$  foi de 327t, enquanto a emissão específica correspondeu a  $0,0324 \text{ tCO}_2/\text{tCaO}$ . Esse valor equivale a

aproximadamente 4% da quantidade de CO<sub>2</sub> emitida na produção de uma tonelada de CaO, inerente ao próprio processo, ou seja, considerando a parcela estequiométrica, e poderia ser evitado por meio da aplicação do modelo em escala industrial.

Observa-se que a maior fração da mitigação potencial de CO<sub>2</sub> está associada às emissões de combustível, responsável por 313tCO<sub>2</sub>, aproximadamente 96% das emissões totais. Tal resultado corrobora a tendência do setor pela busca de combustíveis alternativos para o forno de cal como estratégia de descarbonização da indústria.

Além da mitigação de emissões de CO<sub>2</sub>, a eliminação das condições de sobrequeima resulta em uma economia direta de combustível. Os resultados indicam um potencial de economia de 5.589 GJ de energia térmica. Tratando-se de termos específicos, o período apresentou um potencial de economia de 0,553GJ/tCaO, o que é cerca de 17% da energia necessária para a produção de uma tonelada de óxido de cálcio. Tal valor corresponde a uma economia aproximada de US\$ 105.946,00 em custos de gás natural.

Quando combinado ao valor potencial de créditos de carbono associados à redução de emissões, o benefício econômico total do controle da sobrequeima poderia alcançar US\$107.216,00 no período analisado de 22 meses, evidenciando que a implementação do sensor virtual apresenta vantagens simultaneamente ambientais e econômicas. Anualmente, este valor representa uma economia estimada de US\$56.389,00.

Os valores apresentados devem ser interpretados como estimativas indicativas, uma vez que sua efetiva contabilização depende da elegibilidade do ativo de redução de emissões em mercados específicos, dos padrões de certificação adotados e das condições contratuais estabelecidas. Ainda assim, destaca-se como uma vantagem o fato de que a aplicação do modelo proposto em ambiente industrial não requer modificações no processo produtivo para além do sistema de controle, e nem intervenções estruturais na planta. Por tratar-se de uma ferramenta de suporte à decisão baseada exclusivamente em dados operacionais já disponíveis, sua implementação pode ser realizada de forma incremental e com baixo custo,

minimizando riscos operacionais e facilitando a aceitação por parte das equipes técnicas, o que reforça o seu potencial como solução prática e escalável para a indústria de celulose.

## 7 CONCLUSÕES

O controle de carbonato residual em fornos de cal é um fator relevante para a eficiência energética, a estabilidade operacional, a sustentabilidade ambiental e a qualidade do ciclo de recuperação química na indústria de celulose. Portanto, o desenvolvimento de um modelo de predição do teor de carbonato residual, a partir de dados históricos de processo, conforme proposto neste trabalho, contribui para uma melhor compreensão e monitoramento deste parâmetro crítico.

Foi empregada uma metodologia baseada na análise de dados industriais, envolvendo etapas de pré-processamento, seleção de variáveis e aplicação de técnicas de modelagem preditiva. O desenvolvimento do modelo preditivo baseou-se em nove variáveis operacionais típicas de fornos de cal, incluindo energia específica, temperaturas do leito, do ar primário e do ar secundário, temperatura dos gases de saída, velocidade do ventilador, oxigênio residual, vazão de lama para o filtro e concentração de CO na chaminé. A abordagem empregada utilizou exclusivamente dados *online* de processo, permitindo sua atuação como um sensor virtual. A etapa de pré-processamento a que os dados foram submetidos, principalmente a suavização com o filtro de *Savitzky-Golay*, foi essencial para que apresentassem qualidade suficiente para a obtenção de um modelo representativo e preciso, sem comprometer as informações associadas ao comportamento e à variabilidade do processo de operação do forno.

De modo geral, os modelos baseados em árvores de decisão apresentaram desempenho superior quando comparados aos modelos de redes neurais artificiais. Os resultados indicaram a importância de considerar o tempo de residência do forno de cal, uma vez que os modelos baseados em uma abordagem pontual apresentaram desempenho inferior ao de aqueles que incorporaram um tempo de residência médio de 2,5 horas.

Nesse contexto, as redes neurais artificiais do tipo *Perceptron* Multicamadas não se mostraram capazes de capturar adequadamente as tendências do carbonato residual. O melhor desempenho obtido por MLP apresentou coeficiente de correlação de  $0,45 \pm 0,06$ , erro absoluto médio (MAE) de  $0,56 \pm 0,02\%$  e raiz do erro quadrático médio

(RMSE) de  $0,72 \pm 0,02\%$ , além de elevada dispersão dos resíduos, indicando limitações significativas do modelo e amplo potencial de melhoria.

Em contraste, o algoritmo *Random Forest* resultou em um modelo com desempenho satisfatório, alcançando coeficiente de correlação de  $0,80 \pm 0,01$ , MAE de  $0,36 \pm 0,01\%$  e RMSE de  $0,48 \pm 0,01\%$ . Ainda assim, esse desempenho foi superado pelo algoritmo *XGBoost*, que apresentou os melhores resultados na predição da porcentagem de carbonato residual no produto de calcinação do forno de cal de uma indústria de celulose brasileira. Para esse modelo, obteve-se coeficiente de correlação de  $0,83 \pm 0,01$ , MAE de  $0,34 \pm 0,01\%$ , valor considerado aceitável para aplicações industriais, e RMSE de  $0,45 \pm 0,01\%$ . Adicionalmente, 53% dos resíduos situaram-se no intervalo entre  $[-0,25\%$  e  $+0,25\%]$ , enquanto 78% permaneceram entre  $[-0,5\%$  e  $+0,5\%]$ , o que sugere uma maior precisão e consistência das previsões.

Após a correção de vieses e a construção dos intervalos de confiança, o intervalo correspondente a um desvio padrão abrangeu 73% dos dados reais, sem aplicação de suavização, enquanto o intervalo de dois desvios padrão contemplou 94% dos dados. Tais índices indicam que o modelo corrigido poderia complementar as análises laboratoriais de rotina, por ser capaz de fornecer informações sobre a tendência das condições de operação do forno em tempo real.

O uso de apenas nove variáveis de processo como entrada é um ponto positivo do modelo, pois facilita sua interpretabilidade e também replicação em outras fábricas. A análise de importância das *features* indicou que CO na chaminé, temperatura do ar secundário e velocidade do ventilador apresentaram maior influência para a predição de carbonato residual. Tais variáveis estão diretamente relacionadas à eficiência de combustão, à troca térmica no equipamento e à cinética da reação.

O trabalho foi complementado por uma análise econômica e ambiental simplificada, com o objetivo de fornecer uma estimativa preliminar dos potenciais impactos da aplicação do modelo. Considerando um cenário otimista hipotético, no qual o uso do modelo como sensor virtual contribuiria para a eliminação de casos de supercalcinação, estimou-se um potencial de redução de emissões de CO<sub>2</sub> da ordem de  $0,0324 \text{ tCO}_2/\text{tCaO}$ , sendo aproximadamente 96% desse valor associado à redução do consumo de gás natural.

Adicionalmente, os resultados indicaram um potencial de economia de energia térmica da ordem de 5.589 GJ, equivalente a 0,553 GJ/tCaO, o que corresponde a cerca de 17% da energia necessária para a produção de uma tonelada de óxido de cálcio. Do ponto de vista econômico, o benefício total associado ao controle da sobrequeima poderia alcançar aproximadamente US\$ 107.216,00 no período analisado, o que é equivalente a uma economia estimada de US\$ 56.389,00 em termos anuais. Esses resultados sugerem que a implementação do sensor virtual apresenta potencial para gerar benefícios simultaneamente ambientais e econômicos, devendo, entretanto, ser confirmados por meio de estudos adicionais e validação em ambiente operacional.

De forma geral, os resultados obtidos demonstram que a abordagem baseada em dados proposta neste trabalho apresenta potencial para apoiar o monitoramento e o controle do carbonato residual em fornos de cal da indústria de celulose. Ao integrar técnicas de modelagem preditiva a dados industriais reais, o estudo contribui para o avanço do conhecimento na aplicação de sensores virtuais em processos de recuperação química. Embora as estimativas de desempenho e de benefícios ambientais e econômicos estejam condicionadas às hipóteses adotadas e às limitações dos dados analisados, os resultados indicam que a metodologia desenvolvida constitui uma base consistente para estudos futuros e para eventuais aplicações industriais, e reforça a relevância do uso de ferramentas digitais no aprimoramento da eficiência e da sustentabilidade do setor de celulose.

## 8 RECOMENDAÇÕES

Atingidos os objetivos do trabalho, algumas sugestões são levantadas para estudos futuros.

Inicialmente, recomenda-se o treinamento do modelo preditivo utilizando intervalos de tempo menores, respeitando a máxima resolução de amostragem disponível no ambiente industrial. A utilização de dados com maior frequência temporal pode contribuir para uma melhor captura da dinâmica do processo e para o aumento da sensibilidade do modelo na detecção de variações operacionais e anomalias.

Outra etapa relevante consiste em um teste do modelo em ambiente industrial, permitindo sua avaliação em tempo real e sob condições reais de operação. Essa aplicação possibilitaria analisar o desempenho do sensor virtual de forma contínua, bem como sua integração com sistemas de supervisão e controle, fornecendo informações importantes para sua utilização efetiva como ferramenta de apoio à tomada de decisão.

Adicionalmente, sugere-se a validação da metodologia a partir de dados provenientes de outros fornos de cal, pertencentes a diferentes unidades e indústrias de celulose. Essa análise permitiria avaliar a capacidade de generalização do modelo e identificar a necessidade de ajustes específicos em função de diferentes configurações de equipamentos, matérias-primas e condições operacionais.

No que se refere à análise econômica e ambiental, recomenda-se a realização de avaliações baseadas em dados reais após a implementação do modelo em escala industrial, sem a adoção de simplificações. Isso inclui, por exemplo, o uso de preços reais de gás natural e de crédito de carbono dia a dia, a consideração dos intervalos reais de amostragem para o cálculo de balanços de massa e energia, bem como a quantificação direta das emissões associadas à operação do forno, considerando as premissas de padrões reconhecidos no mercado de carbono.

Por fim, propõe-se a aplicação da metodologia desenvolvida em fornos de outros segmentos industriais que apresentem similaridades com o forno de cal da indústria de celulose, como os fornos utilizados na própria indústria de cal e na indústria de

cimento. Essa extensão permitiria avaliar a adaptabilidade da abordagem a diferentes contextos industriais, ampliando seu potencial de aplicação e contribuição científica.

## REFERÊNCIAS

- ADAMS, T. Lime kiln principles and operations. Atlanta: TAPPI Press, 2008.
- AHMADVAND, S.; SOWLATI, T. The attractiveness of syngas production from forest-based biomass for pulp mills considering carbon pricing and government regulations. **Renewable Energy Focus**, v. 45, p. 287-306, 2023. DOI: <https://doi.org/10.1016/j.ref.2023.04.007>.
- AKBAR, N. A.; DOLLAH, Z.; MOHTAR, Z. A.; ISMAIL, B. N. Effectiveness of innovative teaching and learning strategies for engineering students' performance. **Jurnal kejuruteraan**, v. 37, n. 4, p. 1623–1636, 2025. DOI: [https://doi.org/10.17576/jkukm-2025-37\(4\)-02](https://doi.org/10.17576/jkukm-2025-37(4)-02).
- ALVES, E. D.; PINHEIRO, O. S.; COSTA, A. O. S.; COSTA JUNIOR, E. F. Estudo do processo de obtenção de celulose Kraft com ênfase no forno de cal. **Revista Liberato**, Novo Hamburgo, v. 16, n. 26, p. 101-220, 2015.
- ANUPAM, K.; ANSHU; SHRUTIKONA, D.; GOLEY, P.; DIXIT, A. Box-Behnken design and artificial neural network guided insights on viscosity reduction and maximum swollen volume of heat-treated sugarcane bagasse based kraft black liquor. **Applied Thermal Engineering**, v. 278, n. 127244, p. 127244, 2025. DOI: <https://doi.org/10.1016/j.applthermaleng.2025.127244>
- ARGYROPOULOS, D. *et al.* Kraft lignin: A valuable, sustainable resource, opportunities and challenges. **ChemSusChem**, v.16, n. 23, e202300492, 2023. DOI: <https://doi.org/10.1002/cssc.202300492>
- ARIA, M.; CUCCURULLO, C. bibliometrix : An R-tool for comprehensive science mapping analysis. **Journal of Informetrics**, v. 11, n. 4, p. 959–975, 2017. DOI: <https://doi.org/10.1016/j.joi.2017.08.007>
- ASSUMPÇÃO, R. M.; PINHO, M. R. R.; CAHEN, R.; PHILIPP, P. Polpação química. In: CARVALHO, M. R., & VARALTA, V. T. T. (ed.). **Celulose e Papel: Tecnologia de fabricação da pasta celulósica 2ª ed.** São Paulo: Escola SENAI e IPT, 1988, p. 169-320.
- AZARNIOUCH, M. K.; PHILP, R. J. Wood residues as fuel source for lime kilns - Phase I. Pilot kiln trials. **Pulp and Paper Canada**, v. 84, n. 1, p. 29–34, 1983.
- BARQUIN, J. M. G.; MONSARRO, J. Petrocoke saves money in lime kiln. **Pulp and Paper International**, v. 28, n. 6, p. 60–61, 1986.
- BELISÁRIO, A. B; EDBERG, A.;BJÖRK, M.; ALMEIDA, G. M.; VAKKILAINEN E. K. On the diagnosis of a fouling condition in a kraft recovery boiler: combining process knowledge and data-based insights. **TAPPI Journal**, v. 22, n. 3, p. 162–171, 2023. DOI: <https://doi.org/10.32964/TJ22.3.162>

BERG, E.; HERMANSSON, S.; WETTERLING, J.; LINGMAN, P.; NORDENSKJÖLD, R.; ÅKESJÖ, A. Online monitoring of the size distribution of lime nodules in a full-scale operated lime kiln using an in-situ laser triangulation camera. **TAPPI Journal**, v. 23, n. 6, p. 315–326, 2024. DOI: <https://doi.org/10.32964/TJ23.6.315>

BERGLIN, N; VON SCHENCK, A.; VAKKILAINEN, E. Biofuels in lime kilns – Operating experience in the Nordic pulp and paper industry. **TAPPI Journal**, v. 23, n. 10, p. 559 – 576, 2024. DOI: <https://doi.org/10.32964/TJ23.10.559>

BERGMANN, B. E.; RYING, S. Control of deposit formation and related problems in the bleach plant and pulp mill. **TAPPI**, v. 58, n. 4, p. 147–151, 1975.

BOATENG, A.A. 1—The Rotary Kiln Evolution and Phenomenon. In Boateng, A.A., Ed. **Rotary Kilns**, 2nd ed. Boston: Butterworth-Heinemann, 2016; p. 1–11.

BREIMAN, L. Random Forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.

CARBONE, A.; RAYNER, G. M.; YE, J.; DURANDET, Y. Connecting curricula content with career context: the value of engineering industry site visits to students, academics and industry. **European Journal of Engineering Education**, v. 45, n. 6, p. 971–984, 2020. DOI: <https://doi.org/10.1080/03043797.2020.1806787>

CARTER, R.; ROZEK, I. Kiln control with a virtual residual carbonate analyzer. **Pulp and Paper Canada**, v. 101, n. 8, p. 32–34, 2000.

CAVALCANTE, A. L. B. *et al.* O Impacto das Visitas Técnicas na Percepção dos Alunos quanto à sua Formação Acadêmica. **Revista Científica de Pesquisa Aplicada à Engenharia**, v. 7, n. 1, p. 21–26, 2024.

CHEN, T.; GUESTIN, C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, New York, NY, USA. **Proceedings...** New York: Association for Computing Machinery, 2016. p. 785–794. DOI: <https://doi.org/10.1145/2939672.2939785>

CONFEDERAÇÃO NACIONAL DA INDÚSTRIA. Gás natural: uma análise da formação de preços e da competitividade internacional. Brasília: CNI, 2025. 106 p.

CUI, Y.; YU, C.; SHI, D.; WANG, X. Simulation on coal combustion and calcium carbonate decomposition in a 5500 t/d full scale cement calciner. **Applied Thermal Engineering**, v. 235, n. 121299, p. 121299, 2023. DOI: <https://doi.org/10.1016/j.applthermaleng.2023.121299>

DU, S., HUANG, Z., JIN, L., WAN, X. Recent progress in data-driven intelligent modeling and optimization algorithms for industrial processes. **Algorithms**, v. 17, n. 12, p. 569, 2024. DOI: <https://doi.org/10.3390/a17120569>

DYLONG, M; RAHMAN, M. A comparative study of conventional lime kilns and plasma calcination: Techno-economic assessment and decarbonization potential. In:

**Proceedings of the Second SIMS EUROSIM Conference on Modelling and Simulation, SIMS EUROSIM**, 2024. DOI: <https://doi.org/10.3384/ecp212.013>

ECOSYSTEM MARKETPLACE. **2025 State of the Voluntary Carbon Market (SOVCM): Meeting the Moment – Renewing Trust in Carbon Finance**. 2025. Disponível em: <https://www.ecosystemmarketplace.com/publications/2025-state-of-the-voluntary-carbon-market-sovcm/>. Acesso em: 17 dez 2025.

EPE - EMPRESA DE PESQUISA ENERGÉTICA; IEA - AGÊNCIA INTERNACIONAL DE ENERGIA; IBÁ - INDÚSTRIA BRASILEIRA DE ÁRVORES. **A indústria de papel e celulose no Brasil e no mundo: panorama geral**. EPE, 2022. Disponível em: [https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-650/Pulp%20and%20paper\\_EPE%2BIEA\\_Portugu%C3%AAAs\\_2022\\_01\\_25\\_IBA.pdf](https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-650/Pulp%20and%20paper_EPE%2BIEA_Portugu%C3%AAAs_2022_01_25_IBA.pdf). Acesso em: 11 maio 2025.

EPE- Empresa de Pesquisa Energética. **Balço Energético Nacional – Relatório Síntese 2025, Ano base 2024**. 2025. Disponível em: [https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-885/topico-767/BEN\\_S%C3%ADntese\\_2025\\_PT.pdf](https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-885/topico-767/BEN_S%C3%ADntese_2025_PT.pdf). Acesso em 12 Jan. 2026.

FAO – Food and Agriculture Organization of the United Nations, FAOSTAT, 2025. Disponível em: <https://www.fao.org/faostat/en/#data/FO/visualize>. Acesso em 06 Nov. 2025.

FATOBA, T. M.; OYEYEMI, D. A.; OLUYELE, S. A.; ADENIYI, V. O. Python's indispensability in artificial intelligence and machine learning: a review. **International Journal of Engineering and Modern Technology**, v. 11, n. 4, p. 266–308, 2025. DOI: <https://doi.org/10.56201/ijemt.vol.11.no4.2025.pg266.308>

FERNANDEZ, J. R.; TURRADO, S.; ABANADES, J. C. Calcination kinetics of cement raw meals under various CO<sub>2</sub> concentrations. **Reaction Chemistry & Engineering**, v. 4, n. 12, p. 2129–2140, 2019. DOI: <https://doi.org/10.1039/C9RE00361D>

FRANCEY, S.; TRAN, H. Impacts of burning high-sulphur fuels in lime kilns. **J Sci Technol For Prod Process**, v. 2, n. 3, p. 32–37, 2012.

FRANCEY, S.; TRAN, H.; BERGLIN, N. Global survey on lime kiln operation, energy consumption, and alternative fuel usage. **TAPPI Journal**, v. 10, n. 8, p. 19–26, 2011.

GE, Z. Q., SONG, Z. H., DING, S. X., HUANG, B. Data mining and analytics in the process industry: The role of machine learning. **IEEE Access: Practical Innovations, Open Solutions**, v. 5, p. 20590–20616, 2017. DOI: <https://doi.org/10.1109/ACCESS.2017.2756872>

GRINSZTAJN, L.; OYALLON, E.; VAROQUAUX, G. Why do tree-based models still outperform deep learning on tabular data? **Advances in Neural Information Processing Systems**, 2022. DOI: <https://doi.org/10.48550/arXiv.2207.08815>

GUZZO, G. S. Predição da eficiência de calcinação baseado em variáveis de entrada do forno de cal em uma indústria de celulose. 2024. 74 f. **Monografia** (Especialização em Ciência de Dados) – Instituto de Ciências Exatas e Aplicadas, Universidade Federal de Ouro Preto, João Monlevade, 2024.

HAMPEL, F. R.; RONCHETTI, E. M.; ROUSSEEUW, P. J.; STAHEL, W. A. **Robust statistics: the approach based on influence functions**. New York: Wiley, 1986.

HE, Z.; QIAN, J.; LI, J.; HONG, M.; MAN, Y. Data-driven soft sensors of papermaking process and its application to cleaner production with multi-objective optimization. **Journal of Cleaner Production**, v. 372, p. 133803, 2022. DOI: <https://doi.org/10.1016/j.jclepro.2022.133803>

HODGE, V. J.; AUSTIN, J. A survey of outlier detection methodologies. **Artificial Intelligence Review**, v. 22, n. 2, p. 85–126, out. 2004. DOI: <https://doi.org/10.1023/B:AIRE.0000045502.10941.a9>

HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural Networks: The Official Journal of the International Neural Network Society**, v. 2, n. 5, p. 359–366, 1989.

Hruška, M.; Variny, M.; Haydary, ;., Janošovský, J. Sulfur recovery from syngas in pulp mills with integrated black liquor gasification. **Forests**, v. 11, n. 11, p. 1173, 2020. DOI: <https://doi.org/10.3390/f11111173>

HSIEH, C. L. Evaluating the energy performance of a lime kiln. **TAPPI Journal**, v. 66, n. 9, p. 77–79, 1983.

HUANG, C. W.; BARON, L.; BALAZINSKI, M.; ACHICHE, S. Comprehensive model optimization in pulp quality prediction: a machine learning approach. **PeerJ Preprints**, v. 5, p.e2749, 2017.

IBÁ - Indústria Brasileira de Árvores. **Relatório Anual IBÁ 2025**, 2025. Disponível em: <https://iba.org/wp-content/uploads/2025/10/relatorioAnual2025.pdf>. Acesso em 06 Nov. 2025.

INDIAN PAPER MANUFACTURERS ASSOCIATION. India remains fastest-growing paper market in world with 8.2% rise: report. Business Standard, 4 abr. 2024. Disponível em: [https://www.business-standard.com/industry/news/india-remains-fastest-growing-paper-market-in-world-with-8-2-rise-report-124090400785\\_1.html](https://www.business-standard.com/industry/news/india-remains-fastest-growing-paper-market-in-world-with-8-2-rise-report-124090400785_1.html). Acesso em: 11 abr. 2025.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC). 2006 IPCC Guidelines for National Greenhouse Gas Inventories – Volume 2: Energy. Hayama, Japan: Institute for Global Environmental Strategies (IGES) for the IPCC, 2006. Disponível em: <https://www.ipcc-nggip.iges.or.jp/public/2006gl/>. Acesso em: 26 Fev 2026.

JÄRVENSIVU, M.; KIVIVASARA, J.; SAARI, K. Field survey of TRS emissions from a lime kiln. **Pulp and Paper Canada**, v. 100, n. 11, p. 28–31, 1999.

JÄRVENSIVU, M.; SAARI, K.; JÄMSÄ-JOUNELA, S. L. Intelligent control system of an industrial lime kiln process. **Control Engineering Practice**, v. 9, n. 6, p. 589-606, 2001. DOI: [https://doi.org/10.1016/S0967-0661\(01\)00017-X](https://doi.org/10.1016/S0967-0661(01)00017-X)

JUNEJA, J.; SUNORI, S.; SHARMA, A.; SHARMA, A.; JOSHI, V. Modeling, Control and Instrumentation of Lime Kiln Process: A Review. In: 2020 International Conference on Advances in Computing, Communication & Materials (ICACCM), 2020, India. **Proceedings** [...], Dehradun, 2020, p. 399-403. DOI: <https://doi.org/10.1109/ICACCM50413.2020.9212948>

KARLSTRÖM, A.; HILL, J.; JOHANSSON, L. Data-driven soft sensors in pulp refining processes using artificial neural networks. **BioResources**, v. 19, n. 1, p. 1030–1057, 2024. DOI: <https://doi.org/10.15376/biores.19.1.1030-1057>

KHADIVI, M.; SOWLATI, T. Biomass gasification investment: a multi-criteria decision considering uncertain conditions. **Biomass Conversion and Biorefinery**, v. 14, n. 3, p. 4211-4243, 2024. DOI: <https://doi.org/10.1007/s13399-022-02700-0>

KLUYVER, T. *et al.* Jupyter Notebooks – a publishing format for reproducible computational workflows. In: **LOIZIDES, Fernando; SCHMIDT, Birgit (eds.) Positioning and power in academic publishing: players, agents and agendas.** Göttingen: IOS Press, 2016. p. 87–90. DOI: <https://doi.org/10.3233/978-1-61499-649-1-87>

KUPARINEN, K.; VAKKILAINEN, E. Green pulp mill: renewable alternatives to fossil fuels in lime kiln operations. **BioResources**, v. 12, n. 2, p. 4031–4048, 2017. DOI: <https://doi.org/10.15376/biores.12.2.4031-4048>

KUPARINEN, K.; VAKKILAINEN, E.; TYNJÄLÄ, T. Biomass-based carbon capture and utilization in kraft pulp mills. *Mitigation and Adaptation Strategies for Global Change*, v. 24, n. 7, p. 1213–1230, 2019. <https://doi.org/10.1007/s11027-018-9833-9>

LEE, H.J.; PARK, J.; AND LEE, S.H. Numerical investigation on influence of fan speed and swirling gas injection on thermal-flow characteristics in nitrocarburizing furnace. **Materials Transactions**, v. 58, n. 9, p. 1322–1328, 2017. DOI: <https://doi.org/10.2320/matertrans.M2017148>

LEFVERT, A.; GRÖNKVIST, S.. Smarter ways to capture carbon dioxide – exploring alternatives for small to medium-scale carbon capture in Kraft pulp mills. **International Journal of Greenhouse Gas Control**, v. 127, n. 103934, p. 103934, 2023. DOI: <https://doi.org/10.1016/j.ijggc.2023.103934>

LIANHUA, H.; XINPING, I.; WEI, T.; QINGLI, L. The application of fuzzy control in pulp washing process. In: INTERNATIONAL CONFERENCE ON MEASURING TECHNOLOGY AND MECHATRONICS AUTOMATION, 3., 2011, Shanghai, China. **Proceedings** [...], 2011. p. 686–688. DOI: <https://doi.org/10.1109/ICMTMA.2011.743>

LINDBLOM, J.; THELIANDER, H. Influence of carbon dioxide on ring and ball formation in a pilot scale rotary kiln. **Doktorsavhandlingar Vid Chalmers Tekniska Högskola**, v. 1554, p. 1–18, 1999.

LIPIÄINEN, S.; VAKKILAINEN, E. Role of the Finnish forest industry in mitigating global change: energy use and greenhouse gas emissions towards 2035. **Mitigation and Adaptation Strategies for Global Change**, v. 26, n. 2, 2021. DOI: <https://doi.org/10.1007/s11027-021-09946-5>

LIU, F. T.; TING, K. M.; ZHOU, Z. Isolation Forest. In: **Proceedings of the Eighth IEEE International Conference on Data Mining (ICDM 2008)**, Pisa, Italy, 2008. p. 413–422. DOI: <https://doi.org/10.1109/ICDM.2008.17>

LIU, S.; CHENG, H. Manufacturing Process Optimization in the Process Industry. **International Journal of Information Technology and Web Engineering**, v. 19, n. 1, 2024. DOI: <https://doi.org/10.4018/IJITWE.338998>

LIU, X.; HONG, J.; ZHANG, M.; ZHOU, L. Interpreting the relationship between properties of wood and pulping & paper via machine learning algorithms combined with SHAP analysis. **Nordic Pulp & Paper Research Journal**, v. 40, n. 1, p. 149–160, 2025. DOI: <https://doi.org/10.1515/npprj-2024-0066>

MANNING, R., HANSON, G.M., HART, P.W. The use of parametric analysis to identify opportunities for improved operations of lime kilns. **Peers Conference 2017 Maximizing Success Through Innovation**, 2017-November, pp. 460–475

MANNING, R.; COOPER, S.; MACFADYEN, J. Petcoke firing in lime recovery kilns becomes option as energy costs rise. **Pulp & Paper**, v. 77, n. 12, p. 30–33, 2003.

MARTINEZ, C. M.; SERMYAGINA, E.; SAARI, J.; RAMOS, V. F.; VAKKILAINEN, E.; CARDOSO, M.; ROCHA, E. Fast oxidative pyrolysis of eucalyptus wood residues to replace fossil oil in pulp industry. **Energy**, v. 263, n. 126076, p. 126076, 2023. DOI: <https://doi.org/10.1016/j.energy.2022.126076>

MATEOS-ESPEJEL, E; RADIOTIS, T.; JEEMA, N. Implications of converting a kraft pulp mill to a dissolving pulp operation with a hemicellulose extraction stage. **TAPPI Journal**, v. 12, n. 2, p. 29-38, 2013.

MCCUBBIN, N. Alternatives to fossil fuel for the lime kiln. **Pulp and Paper Canada**, v. 86, n. 9, p. 85–88, 1985.

MICHNIEWICZ, M.; JANIGA, M. The technique for sodium carbonate autocaustization with sodium metaborate. **Przegląd Papierniczy**, v. 66, n. 5, p. 263–268, 2010.

MOBARAKEH, M.; SILVA, M.; KIENBERGER, T.. Pulp and paper industry: Decarbonisation technology assessment to reach CO2 neutral emissions—an Austrian case study. **Energies**, v. 14, n. 4, p. 1161, 2021. DOI: <https://doi.org/10.3390/en14041161>

MOHANAN G, TRAN H, BUSSMAN M, MANNING R. Effect of ring formation on burner flame stability in lime kilns. *TAPPI Journal*, v. 17, n. 05, p. 285–293, 2018.. DOI: <https://doi.org/10.32964/tj17.05.285>

MOLNAR, Christoph. Interpretability — Interpretable Machine Learning. Disponível em: <https://christophm.github.io/interpretable-ml-book/interpretability.htm>. Acesso em: 25 ago. 2025.

MONTGOMERY, D. C. **Introduction to statistical quality control**. 8. ed. Nashville, TN, USA: John Wiley & Sons, 2024.

NEITZEL, I.; KRECZKIUSKI, R. Simulação fluidodinâmica de um digestor contínuo para produção de celulose. In: COMSOL CONFERENCE 2015, Curitiba. **Anais [...]**. Telemaco Borba: FATEB, 2015. Disponível em: <https://www.comsol.com/paper/simula%C3%A7%C3%A3o-fluidodin%C3%A2mica-de-um-digestor-cont%C3%ADnuo-para-produ%C3%A7%C3%A3o-de-celulose-27002>. Acesso em: 18 maio 2025

NEWHART, K. B.; HERING, A. S.; CATH, T. Y. Data science tools to enable decarbonized water and wastewater treatment systems. In: **Pathways to Water Sector Decarbonization, Carbon Capture and Utilization**. IWA Publishing, 2022. p. 275 301.

OIKONOMOU, D.; LEONTARIS, L.; DIMITRIOU, N.; TZOVARAS, D. Time-series forecasting in industrial environments: A performance study and a novel late fusion framework. **IEEE Sensors Journal**, v. 25, n. 4, p. 7681–7697, 2025. DOI: <https://doi.org/10.1109/JSEN.2025.3526362>

OLIVEIRA, G. R.; SILVA, C. L.; OTTO, R. C. C.; GOTHMANN, J. P. S. Visita técnica Como Estratégia de Integração entre Ensino e Prática Profissional: Um relato de Experiência do Centro de Hematologia e Hemoterapia de Santa Catarina – HEMOSC. **Hematology, Transfusion and Cell Therapy**, v. 47, n. 105323, p. 105323, 2025. DOI: <https://doi.org/10.1016/j.htct.2025.105323>

OSMOND, D. R.; TESSIER, P.; SAVOIE, M. Control of an industrial lime kiln operating close to maximum capacity. **TAPPI Journal**, v. 77, n. 2, p. 187–194, 1994.

OTHEN, R.; PADBERG, J.; MÖBITZ, C.; GRIES, T. Digital Transformation in the Paper Industry: Assessing Maturity, Challenges, and Opportunities. **Sustainability**, v. 17, n. 2, p. 770, 2025. DOI: <https://doi.org/10.3390/su17020770>

PENSA, R. G., CROMBACH, A., PEIGNIER, S., RIGOTTI, C. Explaining random forest and XGBoost with shallow decision trees by co-clustering feature importance. **Machine Learning**, v. 114, n. 12, 2025. DOI: <https://doi.org/10.1007/s10994-025-06932-9>

PHILP, R. J.; AZARNIOUCH, M. K. Wood residues as fuel source for lime kilns - Phase II - Pilot kiln trials. **Pulp and Paper Canada**, v. 85, n. 12, p. 114–116, 1984.

PRAKASH, C. B.; MATHUR, K. B.; MURRAY, F. E. Studies on H<sub>2</sub>S emission during calcining. **Pulp and Paper Magazine of Canada**, v. 74, n. 5, p. 99–102, 1973.

RAHMAN, M.; AVELIN, A.; KYPRIANIDIS, K. A Review on the Modeling, Control and Diagnostics of Continuous Pulp Digesters. **Processes**, v. 8, n. 10, 1231, 2020. DOI: <https://doi.org/10.3390/pr8101231>

REN, L.; WANG, H., WANG, Y.; HUANG, K.; WANG, L.; LI, B. Foundation models for the process industry: Challenges and opportunities. **Engineering (Beijing, China)**, v. 52, p. 53–59, 2025. DOI: <https://doi.org/10.1016/j.eng.2025.03.023>

REY, J. R. C.; PIO, D. T.; TARELHO, L. A. C. Biomass direct gasification for electricity generation and natural gas replacement in the lime kilns of the pulp and paper industry: A techno-economic analysis. **Energy**, v. 237, n. 121562, p. 121562, 2021. DOI: <https://doi.org/10.1016/j.energy.2021.121562>

RIBEIRO, B. Prediction of the lime availability on an industrial kiln by neural networks. In: **IEEE World Congress On Computational Intelligence**, 1998, Anchorage. **Proceedings...** Piscataway: IEEE, p. 1987–1991, 1998.

RODRIGUES, J.A.; FARINHA, J. T.; MENDES, M.; MATEUS, R. J. G.; CARDOSO, A. J. M. Comparison of different features and neural networks for predicting industrial paper press condition. **Energies**, v. 15, n. 17, p. 6308, 2022. DOI: <https://doi.org/10.3390/en15176308>

ROGERSON, A., SAVULESCU, L., BERNIER, E., BEDARD, S. Understanding the energy and emission implications of new technologies in a kraft mill: Insights from a CADSIM Plus simulation model. **TAPPI journal**, v. 23, n. 6, p. 336-351, 2024. DOI: <https://doi.org/10.32964//TJ23.6.336>

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, v. 323, n. 6088, p. 533–536, 1986. DOI: <https://doi.org/10.1038/323533a0>

RUSSELL, S.; NORVIG, P. **Artificial intelligence: A modern approach, global edition**. 4. ed. Londres, England: Pearson Education, 2021.

RYAN, J.; BUSSMANN, M.; de MARTINI. CFD and predictive modeling of temperature and calcination in a rotary lime kiln - Potential for steadier kiln operation. **TAPPI Journal**, v. 23, n. 10, p. 545 – 548, 2024. DOI: <https://doi.org/10.32964/TJ23.10.545>

SAIRA, O.; VAUHKONEN, C. Cellulase-assisted refining in a paperboard mill: Avoided emissions from energy savings. A case study of a Finnish paperboard mill. **Cleaner Engineering and Technology**, v. 20, p. 100760, 2024. DOI: <https://doi.org/10.1016/j.clet.2024.100760>

SAREMA, B.; YAN, B. Impact of industrial visits on ECP learners in industrial engineering: Reinforcing the need for a learning factory. In: **Lecture Notes in**

**Networks and Systems.** Cham: Springer Nature Switzerland, 2026. p. 29–36. DOI: [https://doi.org/10.1007/978-3-031-98880-6\\_4](https://doi.org/10.1007/978-3-031-98880-6_4)

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. **Analytical Chemistry**, v. 36, n. 8, p. 1627–1639, 1964.

SCIKIT-LEARN DEVELOPERS. **Ensemble methods — scikit-learn 1.3.2 documentation.** Disponível em: <https://scikit-learn.org/stable/modules/ensemble.html>. Acesso em: 25 ago. 2025b.

SCIKIT-LEARN DEVELOPERS. **Permutation importance — scikit-learn 1.3.2 documentation.** Disponível em: [https://scikit-learn.org/stable/modules/permutation\\_importance.html](https://scikit-learn.org/stable/modules/permutation_importance.html). Acesso em: 25 ago. 2025c.

SCIKIT-LEARN DEVELOPERS. **sklearn.metrics.f1\_score — scikit-learn 1.3.2 documentation.** Disponível em: [https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1\\_score.html](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html). Acesso em: 25 ago. 2025a.

SCOBCZYNSKI, S. F. Reducing energy use through automation: a feasibility study. **PIMA**, v. 66, n. 1, p. 31–34, 1984.

SHAN, W.J., LIU, B., TANG, W. Soft sensing and multi-objective optimization of pulp washing process based on data-driven technology **Journal. of Korea TAPPI**. v. 54, n. 4, p. 57–74, 2022. DOI: <https://doi.org/10.7584/JKTAPPI.2022.08.54.4.57>

SHAVITT, I.; SEGAL, E. Regularization Learning Networks: deep learning for tabular datasets. Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, Montréal, Canada. **Proceedings...** Curran Associates, Inc., 2018. p. 1386–1396.

SOTO, A.; PEREIRA, M. A. Partial replacement of fuel oil by tall oil in the lime kiln. **Celulosa Y Papel**, v. 33, n. 3, p. 22 – 25, 2017.

SOUSA, A.M.; PINTO, I.S.S.; MACHADO, L.; GANDO-FERREIRA, L.; QUINA, M.J. Sustainability of kraft pulp mills: Bleaching technologies and sequences with reduced water use. **Journal of Industrial and Engineering Chemistry**, v. 125, p. 58-70, 2023. DOI: <https://doi.org/10.1016/j.jiec.2023.05.033>

SUÁREZ, V. M. G.; GARCÍA-GONZALO, E.; BAYÓN, R. M.; NIETO, P. J. G.; ANTÓN, J. C. A. Predictive model of gas consumption and air emissions of a lime kiln in a kraft process using the ABC/MARS-based technique. **The International Journal of Advanced Manufacturing Technology**, v. 100, p. 1549–1562, 2018. DOI: <https://doi.org/10.1007/s00170-018-2773-4>

TAILLON, J.; HORVATH, A.; OKSMAN, A. Replacement of fossil fuel with biomass in pulp mill lime kilns. **O Papel**, v. 79, n. 3, p. 85-89, 2018.

TEGUH, N. H.; YULIATI, L.; SISWANTO, E.; DARMADI, D. B. Computational fluid dynamics dataset of secondary air modification in a circulating fluidized bed boiler. **Data in Brief**, v. 52, n. 109931, p. 109931, 2024. DOI: <https://doi.org/10.1016/j.dib.2023.109931>

THAKORE, A. N. Investigation of chemical processes involved in incinerating kraft mill total reduced sulphur gas (TRS) in the lime kilns. **Transactions of the Technical Section of the Canadian Pulp and Paper Association**, v. 5, n. 3, p. 55–58, 1979.

TILLER, P. S.; MIN, B. C.; AHN, H. K.; CARREJO, E.; SANDERS, J.; RIOS, D. C.; PARK, S. Techno-economic analysis of novel lignin extraction from black liquor. **Separation and Purification Technology**, v. 366, p. 132757, 2025. DOI: <https://doi.org/10.1016/j.seppur.2025.132757>

TRAN, H. Kraft recovery operations course proceedings. In: TAPPI KRAFT RECOVERY OPERATIONS COURSE, 2008, St. Petersburg. **Proceedings...** Atlanta: TAPPI Press, 2008.

TRAN, H.; MAO, X.; BARHAM, D. Mechanisms of ring formation in lime kilns. **Journal of Pulp and Paper Science**, v. 19, p. 167–175, 1993.

TRAN, H.; VAKKILAINEN, E. K. The kraft chemical recovery process. Apresentado em: TAPPI Kraft Recovery Short Course, 2008, St. Petersburg, FL., 2008. Disponível em: <https://www.tappi.org/content/events/08kros/manuscripts/1-1.pdf>. Acesso em: 11 maio 2025.

TRAN, H.; VOLLIKS, S.; GAUTHIER, M.; McNEIL, C. Correlation between nodule size and residual carbonate content in lime kilns. **TAPPI Journal**, v. 6, p. n. 4, p. 2007.

TRAUFFER, E. A. Amino scrubbing compounds cut TRS levels with NO<sub>2</sub> reaction. **Pulp and Paper**, v. 69, n. 5, 1995.

TRUNG, T. Advanced online process analyzer for Chemical Recovery and Pulp mill control. **O Papel**, v. 76, n. 1, p. 47- 56, 2015.

TUKEY, J. W. **Exploratory data analysis**. Addison-Wesley, 1977.

URONEN, P.; AURASMAA, H. Modelling and simulation of causticization plant and lime kiln. **Pulp and Paper Canada**, v. 80, n. 6, p. 162–165, 1979.

VAN ECK, N. J.; WALTMAN, L. Software survey: VOSviewer, a computer program for bibliometric mapping. **Scientometrics**, v. 84, n. 2, p. 523–538, 2010. DOI: <https://doi.org/10.1007/s11192-009-0146-3>

VIITALA, R.; MIETTINEN, M.; MARQUEZ, R.; HÄMÄLÄINEN, A.; KARHINEN, A.; BARRIOS, N.; GONZALEZ, R.; PAL, L.; JAMEEL, H.; HOLMBERG, K. Integration of artificial intelligence and sustainable energy management in the pulp and paper industry: A path to decarbonization. **Renewable and Sustainable Energy Reviews**, v 218, 115809, 2025. DOI: <https://doi.org/10.1016/j.rser.2025.115809>

WANG, H; LIANG, Q; HANCOCK, J; KHOSHGOFTAAR, T. Feature selection strategies: a comparative analysis of SHAP-value and importance-based methods. **Journal of Big Data**, v. 11, art. 44, 2024. DOI: <https://doi.org/10.1186/s40537-024-00905-w>

WEI, R. Data analysis with Python: methods, tools and practices — data processing, visualization and prediction based on Python. **Applied and Computational Engineering**, v. 178, p. 112–121, 2025.

WU, R.; CARREJO, E.; REZA, M. D. S.; WOODS, E.; RAZAVI, S.; PARK, S.; LI, F.; SAGUES, W. J. Kinetic assessment of pulp mill-derived lime mud calcination in high CO<sub>2</sub> atmosphere. **Fuel**, v. 373, p .132372, 2024. DOI: <https://doi.org/10.1016/j.fuel.2024.132372>