



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CAMPUS DE RUSSAS**  
**CURSO DE GRADUAÇÃO EM ENGENHARIA MECÂNICA**

**THAYRLA OLIVEIRA ARAÚJO**

**INVESTIGAÇÃO DAS VARIÁVEIS QUE IMPACTAM A PREVISÃO DO  
DIFERENCIAL DE PRESSÃO EM UM MOINHO DE CIMENTO**

**RUSSAS**

**2025**

THAYRLA OLIVEIRA ARAÚJO

INVESTIGAÇÃO DAS VARIÁVEIS QUE IMPACTAM A PREVISÃO DO DIFERENCIAL  
DE PRESSÃO EM UM MOINHO DE CIMENTO

Trabalho de Conclusão de Curso apresentado ao  
Curso de Graduação em Engenharia Mecânica  
do Campus de Russas da Universidade Federal  
do Ceará, como requisito parcial à obtenção do  
grau de bacharel em Engenharia Mecânica.

Orientadora: Profa. Dra. Rosineide Fer-  
nando da Paz

RUSSAS

2025

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

A692i Araujo, Thayrla Oliveira.  
Investigação das Variáveis que Impactam a Previsão do Diferencial de Pressão em um Moinho de Cimento / Thayrla Oliveira Araujo. – 2025.  
69 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Russas, Curso de Engenharia Mecânica, Russas, 2025.  
Orientação: Prof. Dr. Rosineide Fernando da Paz.

1. Diferencial de pressão. 2. Moinho vertical de rolos. 3. Produção de cimento Portland. 4. Random Forest Regressor. I. Título.

CDD 620.1

---

THAYRLA OLIVEIRA ARAÚJO

INVESTIGAÇÃO DAS VARIÁVEIS QUE IMPACTAM A PREVISÃO DO DIFERENCIAL  
DE PRESSÃO EM UM MOINHO DE CIMENTO

Trabalho de Conclusão de Curso apresentado ao  
Curso de Graduação em Engenharia Mecânica  
do Campus de Russas da Universidade Federal  
do Ceará, como requisito parcial à obtenção do  
grau de bacharel em Engenharia Mecânica.

Aprovada em: 27/02/2025

BANCA EXAMINADORA

---

Profa. Dra. Rosineide Fernando da Paz (Orientadora)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Dmontier Pinheiro Aragão Jr.  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Ramon Rudá Brito Medeiros  
Universidade Federal do Ceará (UFC)

Dedico este trabalho a minha mãe, Kátia, que sob muito sol, fez-me chegar até aqui, na sombra.

## AGRADECIMENTOS

Esta dissertação não teria sido possível se desenvolver sem o apoio de diversas pessoas: minha mãe, amigos, orientadores de projeto e colaboradores da Cimentos Apodi Quixeré.

Gostaria de expressar meus sinceros agradecimentos à minha orientadora, Professora Dra. Rosineide Fernando da Paz, por suas palavras de incentivo e empoderamento ao longo dos últimos anos da minha graduação. Suas palavras, "Somos mulheres nas exatas e precisamos ser firmes no que acreditamos", serão sempre um guia em minha trajetória.

Agradeço aos professores orientadores e colegas do Laboratório de Tecnologias Inovadoras (LTI) pela valiosa oportunidade de aprendizado em análise de dados e métodos de aprendizado de máquina. Em especial, ao Luis Antonio Viana Franklin, por sua atenção e apoio nos momentos de insegurança, sempre disposto a me escutar.

Aos colaboradores da Cimentos Apodi, Laís Tamires Oliveira Moraes e Gerson Almeida Ribeiro, agradeço por terem disponibilizado seu tempo e mediado esforços na resolução de dúvidas e incentivos ao projeto.

Expresso uma gratidão imensurável às amigadas que cultivei durante minha graduação, Shirley Naira Carvalho da Silva, Catarina Nogueira Braga, Magna Bocage Irineu, Darlete Vivian Irineu de Sousa, Camila Fernanda Leite de Jesus e Anna Flávia Aragão Oliveira; por terem sido minha principal rede de suporte em Russas e por me proporcionarem momentos de grande afeto.

Agradeço aos colegas do Programa de Ensino Tutorial, os quais convivi intensamente durante os últimos anos, pelo companheirismo e pela troca de experiências que me permitiram crescer não somente como pessoa, mas também como formanda. Em especial ao Gabriel Novais Lima por ter me apoiado e acolhido em circunstâncias adversas.

Às minhas amigas de infância, Francisca Cleisla Carvalho de Sousa e Camila Sampaio Melo, pela amizade incondicional, pelo apoio constante durante todo o período de comprometimento a este trabalho e pelo acolhimento em seus lares, tratando-me como parte da família nos momentos em que mais necessitei.

Ao Janoel Batista Soares, por ter me acompanhado ao longo do meu crescimento e por ter me oferecido o suporte e a orientação de um pai.

Por fim, à minha mãe, Katia de Oliveira Silva Martins, por ter sido o maior pilar de inspiração na minha educação e formação, e por jamais ter poupado esforços para me encorajar

a seguir meus sonhos. Não seria a pessoa que sou hoje sem seu amor e dedicação.

“Necessitamos sempre de ambicionar alguma coisa que, alcançada, não nos torna sem ambição.”

(Carlos Drummond de Andrade)

## RESUMO

O cimento, amplamente utilizado na construção civil, tem sua origem nas civilizações antigas e evoluiu ao longo dos séculos, culminando no desenvolvimento do cimento Portland durante a Revolução Industrial. No Brasil, o cimento Portland é o tipo mais produzido e consumido. Conforme dados do Anuário Estatístico do Brasil, a produção de cimento Portland entre 2020 e 2022 alcançou cerca de 114 milhões de toneladas. O moinho de cimento é um equipamento fundamental na indústria cimenteira, responsável pelo processamento das matérias-primas para a obtenção do produto final. O diferencial de pressão do moinho é um parâmetro crítico de monitoramento, uma vez que seu aumento pode indicar instabilidade no equipamento, resultando em uma elevação atípica da vibração de sua estrutura. No contexto industrial, essa instabilidade compromete a qualidade do cimento, reduz a alimentação do equipamento e acarreta maiores custos energéticos, além de impactar negativamente a produtividade. Diante desse cenário, este estudo analisa as variáveis operacionais associadas ao diferencial de pressão em um moinho vertical de rolos, utilizando uma base de dados histórica de uma unidade fabril. Foram aplicadas técnicas estatísticas para o pré-processamento e suavização dos dados, incluindo análise de defasagem temporal por meio de lag e média móvel, além de estudos de correlações de Pearson, Spearman e de Informação Mútua. Posteriormente, utilizou-se aprendizado de máquina, com o modelo Random Forest Regressor, para prever o comportamento do diferencial de pressão. Os resultados indicam que as correlações de Pearson e Spearman corroboraram a relevância das variáveis analisadas, enquanto a Informação Mútua evidenciou interações mais complexas e redundâncias informacionais, sugerindo a possibilidade de simplificação do modelo preditivo. A modelagem apresentou potencial preditivo, porém as métricas de desempenho do modelo indicam a necessidade de ajustes adicionais para melhorar a precisão das previsões.

**Palavras-chave:** diferencial de pressão; moinho vertical de rolos; produção de cimento Portland; Random Forest Regressor

## ABSTRACT

Cement, widely used in construction, has its origins in ancient civilizations and has evolved over the centuries, culminating in the development of Portland cement during the Industrial Revolution. In Brazil, Portland cement is the most produced and consumed type. According to data from the Statistical Yearbook of Brazil, the production of Portland cement between 2020 and 2022 reached approximately 114 million tons. The cement mill is a fundamental piece of equipment in the cement industry, responsible for processing raw materials to obtain the final product. The mill's pressure differential is a critical monitoring parameter, since its increase may indicate instability in the equipment, resulting in an atypical increase in the vibration of its structure. In the industrial context, this instability compromises the quality of the cement, reduces the equipment's feed and leads to higher energy costs, in addition to negatively impacting productivity. In view of this scenario, this study analyzes the operational variables associated with the pressure differential in a vertical roller mill, using a historical database from a manufacturing unit. Statistical techniques were applied for preprocessing and smoothing the data, including time lag analysis using lag and moving average, in addition to Pearson, Spearman and Mutual Information correlation studies. Subsequently, machine learning was used, with the Random Forest Regressor model, to predict the behavior of the pressure differential. The results indicate that the Pearson and Spearman correlations corroborated the relevance of the analyzed variables, while the Mutual Information showed more complex interactions and informational redundancies, suggesting the possibility of simplifying the predictive model. The modeling showed predictive potential, but the model's performance metrics indicate the need for additional adjustments to improve the accuracy of the predictions.

**Keywords:** differential pressure; vertical roller mill; Portland cement production; Random Forest Regressor

## LISTA DE FIGURAS

Figura 1 – Fluxograma das etapas do processo produtivo de cimento portland . . . . .	21
Figura 2 – Limites de composição do cimento Portland . . . . .	23
Figura 3 – Componentes de moinho vertical de rolos . . . . .	24
Figura 4 – Fluxograma da metodologia do estudo. . . . .	31
Figura 5 – Exemplos de gráficos de dispersão de correlação de Pearson . . . . .	33
Figura 6 – Fluxograma das etapas de modelagem. . . . .	36
Figura 7 – Ilustração de cada tipo de partição para cross-validation. . . . .	39
Figura 8 – Gráficos histograma e bloxpot do diferencial de pressão. . . . .	41
Figura 9 – Gráficos histograma e bloxpot da alimentação total. . . . .	42
Figura 10 – Gráficos histograma e bloxpot do <i>dampner 02</i> . . . . .	42
Figura 11 – Gráficos histograma e bloxpot da injeção de água. . . . .	43
Figura 12 – Gráficos histograma e bloxpot da pressão de entrada. . . . .	43
Figura 13 – Gráficos histograma e bloxpot da pressão de moagem. . . . .	44
Figura 14 – Gráficos histograma e bloxpot da temperatura de entrada. . . . .	44
Figura 15 – Gráficos histograma e bloxpot da temperatura de entrada. . . . .	45
Figura 16 – Gráficos histograma e bloxpot da temperatura de saída. . . . .	46
Figura 17 – Gráficos histograma e bloxpot da velocidade do classificador. . . . .	46
Figura 18 – Correlação cruzada de Pearson para diferentes valores de média móvel . . .	50
Figura 19 – Correlação cruzada de Spearman para diferentes valores de média móvel . .	51
Figura 20 – Correlação cruzada de Informação Mútua para diferentes valores de média móvel . . . . .	52
Figura 21 – Matrizes <i>heatmap</i> da correlação de Pearson: (a) com os dados não tratados, (b) com os dados filtrados e (c) com a aplicação das suavizações. . . . .	53
Figura 22 – Matrizes <i>heatmap</i> da correlação de Spearman: (a) com os dados não tratados, (b) com os dados filtrados e (c) com a aplicação das suavizações. . . . .	54
Figura 23 – Matrizes <i>heatmap</i> da correlação de Informação Mútua: (a) com os dados não tratados, (b) com os dados filtrados e (c) com a aplicação das suavizações. . .	56
Figura 24 – Gráfico dos valores observados e preditos do diferencial de pressão nas amostras de teste. . . . .	57
Figura 25 – Análise SHAP das variáveis que influenciam a predição do diferencial de pressão do moinho. . . . .	57

Figura 26 – Gráfico da evolução do erro absoluto médio dos grupos para cada combinação de parâmetros . . . . .	59
Figura 27 – Gráfico da evolução do erro quadrático médio da raiz dos dos grupos para cada combinação de parâmetros . . . . .	60
Figura 28 – Gráfico da evolução do erro absoluto médio escalonado dos grupos para cada combinação de parâmetros . . . . .	60
Figura 29 – Gráfico dos valores observados e preditos do diferencial de pressão nas amostras de teste, com a parametrização ajustada. . . . .	61
Figura 30 – Gráfico dos valores observados e preditos do diferencial de pressão na amostra de teste para os dados com baixa variabilidade de pressão de saída. . . . .	64

## LISTA DE TABELAS

Tabela 1 – Interpretação do coeficiente de correlação de Pearson . . . . .	34
Tabela 2 – Medidas estatísticas sobre as variáveis analisadas . . . . .	40
Tabela 3 – Medidas estatísticas sobre as variáveis após a filtragem . . . . .	47
Tabela 4 – Tabela de valores de lag e janela de média móvel para os valores máximos de correlação de Pearson . . . . .	48
Tabela 5 – Tabela de valores de lag e janela de média móvel para os valores máximos de correlação de Spearman . . . . .	48
Tabela 6 – Tabela de valores de lag e janela de média móvel para os valores máximos da correlação de Informação Mútua . . . . .	48
Tabela 7 – Tabela de valores de lag e janela de média móvel selecionados para suavização de dados . . . . .	53
Tabela 8 – Valores das métricas para teste da predição não parametrizada. . . . .	57
Tabela 9 – Combinação de valores de parâmetros para validação cruzada . . . . .	59
Tabela 10 – Valores das métricas para teste da predição parametrizada. . . . .	61
Tabela 11 – Tabela dos valores de lag e da janela de média móvel para os valores máximos de correlação de Pearson, referentes a dados com baixa variabilidade da pressão de saída. . . . .	62
Tabela 12 – Tabela dos valores de lag e da janela de média móvel para os valores máximos de correlação de Spearman, referentes a dados com baixa variabilidade de pressão de saída. . . . .	62
Tabela 13 – Tabela dos valores de lag e da janela de média móvel para os valores máximos de Informação Mútua, referentes a dados com baixa variabilidade da pressão de saída. . . . .	63
Tabela 14 – Tabela de valores de lag e janela de média móvel para dados com baixa variabilidade de alimentação total . . . . .	63
Tabela 15 – Métricas de desempenho do modelo de predição com baixa variabilidade de pressão de saída . . . . .	63



## LISTA DE SÍMBOLOS

<i>%</i>	Porcentagem
<i>°C</i>	Graus Celsius
<i>h</i>	Hora
<i>m</i>	Metro
<i>mbar</i>	Milibar
<i>rpm</i>	Rotações por minuto
<i>t</i>	Tonelada

## SUMÁRIO

1	<b>INTRODUÇÃO</b>	17
2	<b>OBJETIVOS E JUSTIFICATIVA</b>	19
2.1	Objetivo geral	19
2.2	Objetivos secundários	19
2.3	Justificativa	19
3	<b>FUNDAMENTAÇÃO TEÓRICA</b>	20
3.1	<b>Fabricação de cimento</b>	20
3.1.1	<i>Farinha crua</i>	20
3.1.2	<i>Clínquer</i>	21
3.1.3	<i>Cimento</i>	22
3.2	<b>Variável resposta e variáveis explicativas</b>	24
3.2.1	<i>Diferencial de pressão</i>	24
3.2.2	<i>Pressão de entrada</i>	25
3.2.3	<i>Pressão de saída</i>	25
3.2.4	<i>Temperatura de entrada</i>	25
3.2.5	<i>Temperatura de saída</i>	25
3.2.6	<i>Arraste de gases</i>	25
3.2.7	<i>Recirculação de gases</i>	26
3.2.8	<i>Alimentação total</i>	26
3.2.9	<i>Injeção de água</i>	26
3.2.10	<i>Pressão de moagem</i>	26
3.2.11	<i>Velocidade do classificador</i>	27
4	<b>TRABALHOS RELACIONADOS</b>	28
4.1	<b>Estudo de Variáveis que Afetam a Vibração do Corpo de um Moinho Vertical de Rolos (MENESES, 2023)</b>	28
4.2	<i>Power-draw prediction by random forest based on operating parameters for an industrial ball mill (BELMAJDOUB, 2019)</i>	29
4.3	<i>Efficient machine learning model to predict fineness, in a vertical raw meal of Morocco cement plant (TOHRY S. CHEHREH CHELGANI, 2022)</i>	29
5	<b>METODOLOGIA</b>	31

5.1	Coleta de dados . . . . .	31
5.2	Pré-processamento de dados . . . . .	31
5.2.1	<i>Crítica aos dados</i> . . . . .	31
5.2.2	<i>Média móvel</i> . . . . .	32
5.2.3	<i>Lag</i> . . . . .	32
5.3	Correlações . . . . .	33
5.3.1	<i>Correlação de Pearson</i> . . . . .	33
5.3.2	<i>Correlação de informação mútua</i> . . . . .	34
5.3.3	<i>Correlação de Spearman</i> . . . . .	35
5.4	Aprendizado de máquina . . . . .	36
6	<b>RESULTADOS</b> . . . . .	40
6.1	Análise exploratória de dados . . . . .	40
6.1.1	<i>Análise do diferencial de pressão</i> . . . . .	41
6.1.2	<i>Análise da alimentação total</i> . . . . .	41
6.1.3	<i>Análise do arraste de gases</i> . . . . .	42
6.1.4	<i>Análise da injeção de água</i> . . . . .	42
6.1.5	<i>Análise da pressão de entrada</i> . . . . .	43
6.1.6	<i>Análise da pressão de moagem</i> . . . . .	43
6.1.7	<i>Análise da pressão de saída</i> . . . . .	44
6.1.8	<i>Análise da temperatura de entrada</i> . . . . .	45
6.1.9	<i>Análise da temperatura de saída</i> . . . . .	45
6.1.10	<i>Análise da velocidade do classificador</i> . . . . .	45
6.2	Análise da defasagem de tempo . . . . .	46
6.3	Análise das relações via correlação de Pearson . . . . .	49
6.4	Análise das relações via correlação de Spearman . . . . .	51
6.5	Análise das relações via Informação Mútua . . . . .	52
6.6	Predição com <i>Random Forest Regressor</i> . . . . .	55
6.6.1	<i>Configuração dos parâmetros do modelo</i> . . . . .	58
6.6.2	<i>Predição com baixa variabilidade de pressão de saída</i> . . . . .	59
7	<b>CONSIDERAÇÕES FINAIS</b> . . . . .	65
	<b>REFERÊNCIAS</b> . . . . .	67

## 1 INTRODUÇÃO

A origem do cimento remonta às civilizações antigas há cerca de 4.500 anos atrás. Os egípcios empregavam gesso calcinado como ligante, enquanto os romanos desenvolveram um material mais avançado ao misturar cal com cinzas vulcânicas, conhecidas como pozolana, criando um "cimento romano" utilizado em estruturas como o Panteão. No entanto, foi durante a Revolução Industrial, em 1824, que Joseph Aspdin patenteou o cimento Portland, produzido a partir da calcinação de uma mistura de calcário e argila (ABCP, 2024).

O cimento Portland, base do cimento moderno, evoluiu com melhorias tecnológicas para atender às demandas da construção civil. No Brasil, é o tipo mais produzido e consumido. Segundo o Anuário Estatístico do Brasil, elaborado pelo IBGE (2023), a produção de cimento Portland entre 2020 e 2022 foi de cerca de 114 milhões de toneladas. Sua fabricação envolve a extração de matérias-primas, como calcário e argila, a queima em forno rotativo para formar o clínquer e, por fim, a moagem para obter o cimento.

O moinho de cimento é o equipamento utilizado na indústria de cimento para moer clínquer juntamente com outros componentes como gesso, escória ou aditivo. Existem dois tipos usuais de moinho, o moinho de bolas e o moinho vertical de rolos. As principais diferenças entre eles estão no quesito operacional e eficiência energética. O moinho vertical de rolos é mais eficiente energeticamente, apresentando menor desgaste e necessidade de manutenção, especialmente em grandes produções contínuas.

Além disso, o moinho vertical de rolos proporciona maior controle sobre a finura do produto, resultando em cimento de qualidade superior e homogeneidade. O diferencial de pressão no moinho, parâmetro crucial para assegurar o fluxo contínuo de material, representa a diferença entre as pressões de entrada e saída. O controle adequado desse diferencial indica a eficiência da moagem, mantendo o equilíbrio entre a entrada de matéria-prima e a saída do cimento, prevenindo desproporções no processo.

O aumento do diferencial de pressão além dos parâmetros operacionais ideais em um moinho de cimento pode ocasionar instabilidade no equipamento, o que, conseqüentemente, leva a um aumento anômalo na vibração do corpo do moinho. No contexto operacional, a fim de evitar prejuízos à qualidade do cimento, essa instabilidade resulta na redução da alimentação do equipamento, acarretando um aumento nos custos energéticos em detrimento da queda na produtividade.

Neste contexto, o presente estudo tem como objetivo analisar as variáveis de controle

operacional associadas ao diferencial de pressão em um moinho vertical de rolos. Para alcançar tal objetivo, as análises foram realizadas com o uso de uma base de dados histórica oriunda de uma unidade fabril cimenteira, abrangendo o período de outubro de 2024 a janeiro de 2025. Essas análises envolvem o processo de pré-processamento dos dados, a análise de correlações, a modelagem e a predição da variável diferencial de pressão.

As correlações aplicadas, incluindo Pearson, Spearman e Informação Mútua, visaram compreender as relações entre as variáveis explicativas e a variável resposta, com o intuito de identificar tendências e estabelecer valores de lag (defasagem temporal) e janelas de média móvel, técnicas utilizadas para avaliar o efeito das variáveis ao longo do tempo e atenuar flutuações nos dados. A correlação de Pearson demonstrou maiores valores para a maioria das variáveis. As correlações de Spearman corroboraram os resultados obtidos por Pearson.

A aplicação da Informação Mútua, por sua vez, evidenciou a complexidade das interações, revelando correlações mais robustas entre variáveis inicialmente identificadas com relações fracas. Subsequentemente, foi implementada a modelagem baseada em aprendizado de máquina, utilizando o algoritmo *Random Forest Regressor*. A partir dessa modelagem, configurou-se o comportamento da predição com os parâmetros definidos, evidenciando que o modelo apresentou melhor desempenho quando aplicado a uma seleção de valores com baixa variabilidade da pressão de saída.

## 2 OBJETIVOS E JUSTIFICATIVA

Este capítulo tem como propósito apresentar o objetivo geral, os objetivos específicos e a justificativas que embasaram a condução deste trabalho.

### 2.1 Objetivo geral

O presente trabalho estudou as principais variáveis influentes na variável diferencial de pressão de um moinho vertical de rolos para produção de cimento Portland em uma planta cimenteira.

### 2.2 Objetivos secundários

Objetivando alcançar o objetivo geral deste trabalho, foram delimitados os seguintes objetivos específicos:

- Elaborar uma base de dados tratada para a análise da variável "diferencial de pressão";
- Realizar uma análise estatística exploratória, com foco nas correlações entre as variáveis explicativas e a variável "diferencial de pressão";
- Investigar os efeitos da defasagem temporal sobre a relação entre as variáveis;
- Conduzir uma análise de correlações utilizando os métodos de Pearso e Spearman;
- Determinar e aplicar valores de média móvel adequados para as variáveis explicativas;
- Desenvolver um modelo de predição utilizando técnicas de aprendizado de máquina, por meio do algoritmo *Random Forest Regressor*.

### 2.3 Justificativa

O aumento do diferencial de pressão em um moinho de cimento pode causar instabilidade no equipamento, o que, por sua vez, gera um aumento atípico na vibração do corpo do moinho. No contexto das operações industriais, a fim de preservar a qualidade do cimento, essa instabilidade resulta na redução da alimentação do equipamento, gerando elevados custos energéticos e redução na produtividade.

### 3 FUNDAMENTAÇÃO TEÓRICA

A fundamentação teórica deste estudo tem como objetivo apresentar os principais conceitos relacionados à fabricação de cimento, bem como as variáveis envolvidas no processo de moagem de cimento em moinho vertical de rolos, com o intuito de fornecer o embasamento necessário para o tema em questão.

#### 3.1 Fabricação de cimento

O cimento é um material de construção em pó que quando misturado com água, forma uma pasta que endurece e adquire resistência ao longo do tempo (MEHTA; MONTEIRO, 2006). No Brasil, o tipo de cimento mais amplamente utilizado é o cimento Portland, devido à sua alta resistência mecânica, durabilidade e versatilidade. Sua aplicação está em diversos tipos de estruturas, desde residências até grandes obras de infraestrutura, como pontes, barragens e estradas.

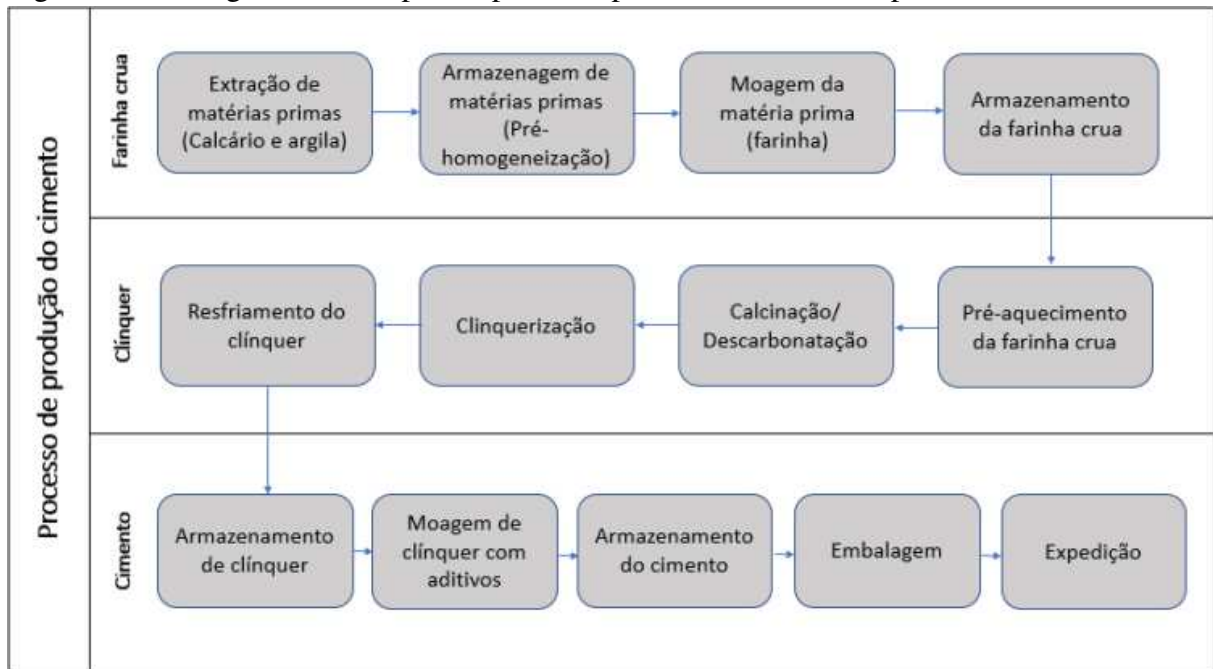
De acordo com Lima (2011), o processo de fabricação de cimento envolve diversas etapas, começando com a extração das matérias-primas, passando pela moagem e homogeneização para formar a "farinha de cru", seguida da clínquerização, etapa na qual os materiais são submetidos a altas temperaturas para a produção do clínquer, e finalizando com a moagem do clínquer com adições para obtenção do cimento. A Figura 1 apresenta um fluxograma do processo produtivo.

##### 3.1.1 *Farinha crua*

As matérias-primas essenciais para a produção de cimento são calcário e argila. O calcário é a principal fonte de carbonato de cálcio, que constitui a base para a produção do clínquer. Por outro lado, a argila é um componente crucial, pois fornece silicatos de alumínio e sílica, fundamentais para equilibrar a composição química do cimento (HEWLETT, 2019). A areia pode ser utilizada em algumas formulações para ajustar a proporção de sílica.

A extração desses minerais é realizada por meio da mineração a céu aberto, com a escavação e transporte de calcário e argila para a planta de produção. Em muitos casos, utiliza-se explosivos para fragmentar grandes blocos de rocha, facilitando o processo de extração. Após esta etapa, as matérias-primas passam pela britagem, onde são reduzidas a tamanhos adequados para o processamento. O controle da granulometria durante a britagem interfere diretamente na

Figura 1 – Fluxograma das etapas do processo produtivo de cimento portland



Fonte: Meneses (2023).

qualidade do clínquer e, por consequência, do cimento produzido (SOUZA *et al.*, 2020).

Em seguida, as matérias-primas passam pela pré-homogeneização, onde calcário e argila são empilhados em camadas alternadas ou armazenados em silos que promovem uma mistura mais homogênea. A aglomeração mineral desse modo é denominada Pilha Chevron. Esse processo é essencial para garantir uma composição química uniforme, influenciando diretamente a qualidade do clínquer. Após o empilhamento, a mistura é transportada para silos de dosagem, que garantem estoque e alimentação constante aos moinhos.

O argical, que se refere à mistura de argila e calcário, constitui a principal matéria-prima empregada na produção da farinha crua ou cru. A preparação dos componentes do cru é realizada por meio da moagem da matéria-prima em moinhos de bolas e moinhos verticais, com o objetivo de atingir uma granulometria adequada (ISAIA, 2017). Em seguida, a farinha é armazenada em silos de homogeneização, com a finalidade de assegurar a consistência de sua composição química e servir como um pulmão para a alimentação do forno rotativo.

### 3.1.2 Clínquer

A produção do clínquer é realizada através da exposição da farinha crua a altas temperaturas, aproximadamente 1450 °C, em fornos rotativos horizontais (COSTA *et al.*, 2013). A primeira reação que ocorre durante esse processo é a calcinação, também conhecida como

descarbonatação. Essa etapa envolve o aquecimento do material até atingir uma temperatura que provoca a decomposição química do carbonato de cálcio, resultando na formação de óxido de cálcio e na liberação de dióxido de carbono.

O material é submetido a um aquecimento adicional até atingir a fusão inicial, momento em que o material aglutina em pelotas, formando grânulos duros e densos, denominados clínquer. Este processo é conhecido como clinquerização. Após a formação, o material é resfriado rapidamente, geralmente com o uso de ar ou água, a fim de interromper o processo químico e preservar suas propriedades. Em seguida, o insumo é armazenado em silos apropriados.

A configuração de um forno giratório consiste em um cilindro longo e levemente inclinado, geralmente fabricado em aço e revestido com material refratário. A inclinação do forno combinada ao movimento lento de rotação é fundamental para permitir que o material deslize lentamente em direção à saída à medida que o cilindro gira. Na parte inferior, há queimadores alimentados por combustíveis que são responsáveis pelo aquecimento interno, gerando calor necessário para as reações químicas que ocorrem durante o processo (NEVILLE, 2011).

### **3.1.3 Cimento**

Os materiais processados no moinho de cimento variam de acordo com o tipo de cimento em produção. A fabricação do Cimento Portland é realizada em conformidade com a norma NBR 16697 da ABNT, intitulada "Cimento Portland – Requisitos". Esta estabelece a classificação dos tipos de cimento, suas propriedades físicas e químicas, requisitos químicos, métodos de ensaio, além das diretrizes para marcação e rotulagem. A Figura 2 apresenta a designação padronizada pela referida norma.

A escória granulada de alto-forno, subproduto da fabricação de ferro-gusa, é utilizada como matéria-prima sustentável na produção de cimento, contribuindo para a redução das emissões de dióxido de carbono (SOUSA *et al.*, 2024). Como material pozolânico, ela reage com o hidróxido de cálcio durante a hidratação do clínquer, formando compostos adicionais que melhoram as propriedades mecânicas e a durabilidade do cimento, além de reduzir o consumo de clínquer.

A adição de gesso tem como objetivo regular o tempo de pega do cimento, uma vez que a hidratação rápida do clínquer resulta em seu endurecimento precoce, prejudicando a manipulação adequada do material. Essas matérias-primas são dosadas e incorporadas ao processo de moagem do cimento, realizado em um moinho.

Figura 2 – Limites de composição do cimento Portland

Designação normalizada		Sigla	Classe de resistência	Sufixo	Clinker + sulfatos de cálcio	Escória granulada de alto-forno	Material pozzolânico	Material carbonático
Cimento Portland comum		CP I	25, 32 ou 40	RS ou BC	95 – 100	0 – 5		
		CP I-S			90 – 94	0	0	6 – 10
Cimento Portland composto com escória granulada de alto-forno		CP I-IE			51 – 94	6 – 34	0	0 – 15
Cimento Portland composto com material pozzolânico		CP II-Z			71 – 94	0	6 – 14	0 – 15
Cimento Portland composto com material carbonático		CP II-F			75 – 89	0	0	11 – 25
Cimento Portland de alto forno		CP III			25 – 65	35 – 75	0	0 – 10
Cimento Portland pozzolânico		CP IV			45 – 85	0	15 – 50	0 – 10
Cimento Portland de alta resistência inicial		CP V <sup>a</sup>	ARI		90 – 100	0	0	0 – 10
Cimento Portland branco	Estrutural	CPB	25, 32 ou 40		75 – 100	–	–	0 – 25
	Não estrutural		–	–	50 – 74	–	–	26 – 50

<sup>a</sup> No caso de cimento Portland de alta resistência inicial resistente a sulfatos (CP V-ARI RS), podem ser adicionadas escórias granuladas de alto-forno ou materiais pozzolânicos.

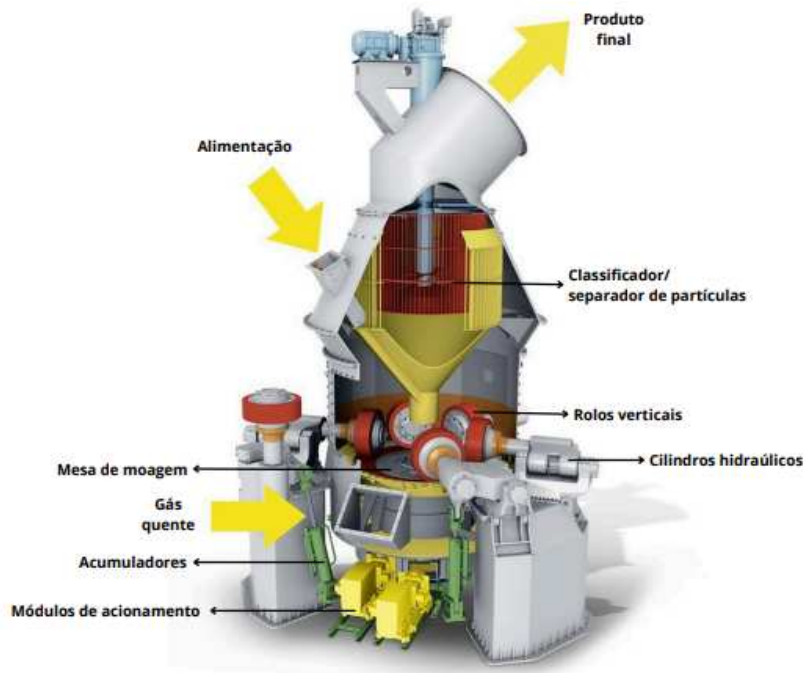
Fonte: ABNT (2018).

A Figura 3 apresenta os principais componentes do moinho estudado. A alimentação do equipamento é realizada por meio de uma calha, que deposita as dosagens das matérias-primas que compõem a receita do cimento no interior do equipamento. A mistura é moída entre os rolos e a mesa de moagem giratória. Os braços de alavanca, conectados a um sistema hidropneumático que atua sobre os rolos, são responsáveis por aplicar a pressão de moagem sobre a superfície da mesa (REICHERT *et al.*, 2015). Durante esse processo, a água é injetada no interior do moinho para controlar a temperatura gerada e facilitar a moagem.

O equipamento apresenta quatro rolos, cada um conectado a cilindros hidráulicos de forma individual. Assim, a pressão de moagem corresponde, tecnicamente, à pressão do óleo na parte superior do pistão. Nesse sentido, as pressões aplicadas nos pares de rolos são equivalentes. O coque, um combustível com elevado poder calorífico, proporciona a energia térmica necessária para a evaporação da água contida nas matérias-primas. Este aspecto é relevante, uma vez que materiais úmidos tendem a se aglomerar, dificultando o processo de moagem.

O material selecionado é direcionado para o classificador, localizado na parte superior do moinho, onde é segregado conforme sua finura. No interior do classificador, uma força centrífuga, assistida por um fluxo de ar, eleva as partículas moídas. As partículas mais leves e finas são conduzidas para a saída e, por meio de filtros de manga, são encaminhadas para os

Figura 3 – Componentes de moinho vertical de rolos



Fonte: Pfeirer (1940).

silos destinados ao armazenamento do cimento. Já o material mais grosso, que não atende à granulometria estabelecida, é recirculado para a zona de moagem para um novo processo de trituração.

### 3.2 Variável resposta e variáveis explicativas

Esta seção aborda o diferencial de pressão, a partir deste ponto denominado como variável resposta, assim como as variáveis que o influenciam, referidas como variáveis explicativas em virtude de sua influência sobre esta variável.

#### 3.2.1 Diferencial de pressão

O diferencial de pressão no moinho refere-se à diferença entre a pressão de entrada e a pressão de saída do equipamento, sendo expresso em milibar *mbar*. Este parâmetro é fundamental para garantir um fluxo constante de material no interior do equipamento, evitando desproporcionalidades entre a entrada de compostos e a saída de cimento.

A pressão de entrada é variável até que a exaustão e a recirculação de gases sejam devidamente ajustadas; a partir desse ponto, ela se estabiliza em um valor fixo. Em contrapartida, a pressão de saída tende ao negativo à medida que a alimentação do moinho aumenta, uma vez que está relacionada ao duto de saída do moinho.

### **3.2.2 Pressão de entrada**

A pressão de entrada do moinho refere-se à pressão do ar ou dos gases introduzidos no sistema do moinho no ponto de entrada. Está intimamente ligada ao ajuste do arraste e recirculação dos gases introduzidos e produzidos durante o processo de moagem.

### **3.2.3 Pressão de saída**

A pressão de saída está diretamente associada ao fluxo de materiais e gases que deixam o sistema do moinho. Essa variável é influenciada pela velocidade de classificação dos materiais e pela configuração do *damper*. Um menor valor de velocidade do classificador resulta em uma pressão de saída mais elevada, indicando que uma maior quantidade de material está sendo descarregada do moinho.

### **3.2.4 Temperatura de entrada**

A temperatura de entrada do moinho diz respeito à temperatura dos gases ou do ar introduzido durante o processo de moagem. Está variável é influenciada pela utilização de coque como fonte de calor e energia. A combustão do coque gera gases que auxiliam na remoção da umidade das matérias-primas. Com o aumento da temperatura, os gases se expandem conforme o princípio dos gases ideais, resultando no aumento do diferencial de pressão do moinho.

### **3.2.5 Temperatura de saída**

A temperatura de saída do moinho diz respeito ao fluxo de gases e materiais que sai do moinho após o processo de moagem. Ela reflete o estado térmico final do sistema.

### **3.2.6 Arraste de gases**

O controle do fluxo gasoso é realizado por meio da abertura e fechamento de um dispositivo conhecido como *damper* DA01. Este dispositivo é avaliado em porcentagem, onde 100 representa a abertura total. Um aumento anômalo na temperatura interna do moinho provoca a expansão dos gases, o que dificulta a triagem gasosa e eleva a pressão interna do equipamento. Essa condição agrava o diferencial de pressão do sistema.

### **3.2.7 Recirculação de gases**

A recirculação de gases é gerenciada por um *damper* adicional, denominado DA02. Essa estrutura é responsável pelo reaproveitamento dos gases, contribuindo para a redução do consumo térmico. A mesma relação da temperatura com arraste de gases é aplicado para esta variável.

### **3.2.8 Alimentação total**

A alimentação total refere-se ao conjunto de dosagens das matérias-primas empregadas na moagem de cimento, abrangendo calcário, gesso, escória e clínquer. Esses componentes são dosados conforme a composição do tipo de cimento analisado e apresentado no tópico anterior. Um aumento atípico na alimentação do moinho gera dificuldades na exaustão dos gases, ultrapassando o limite máximo de pressão operacional.

Como consequência, o fluxo de transporte do material é comprometido, o que resulta no retorno do material moído à mesa de moagem, o qual não é classificado pelo separador e permanece no interior do moinho. Nesse contexto, a pressão de saída torna-se progressivamente mais negativa devido à restrição do material no duto de saída, ocasionando um aumento substancial no diferencial de pressão.

### **3.2.9 Injeção de água**

A injeção de água é responsável por acomodar regularmente o material na mesa, impedindo a sobreposição de partículas e contribuindo para uma moagem mais uniforme. A proporção de injeção de água no sistema varia de 3% a 5%, de acordo com a alimentação total. Um aumento não planejado na injeção de água pode resultar no aglomeramento do material, formando lama e aumentando a resistência ao fluxo dos materiais no moinho, o que leva a um incremento no diferencial de pressão.

### **3.2.10 Pressão de moagem**

A pressão de moagem é a força exercida pelos rolos sobre o leito de material na mesa de moagem. Essa força é fundamental para a eficácia do processo de moagem, pois influencia tanto a taxa de desgaste do material quanto a eficiência na redução do tamanho das partículas. Uma redução na pressão de moagem pode resultar em um decréscimo no diferencial de pressão,

o que indica que o material permanece por um período mais prolongado dentro do moinho.

### **3.2.11 Velocidade do classificador**

O classificador é uma estrutura posicionada na parte superior do moinho, encarregada de selecionar a finura do produto com base na sua velocidade de rotações por minuto (rpm). Um aumento desta velocidade resulta em critérios mais rigorosos para a seleção de partículas, o que resulta em uma maior retenção de material. Esse fenômeno prolonga o tempo de moagem e, conseqüentemente, eleva o diferencial de pressão no sistema.

Um aumento no diferencial de pressão de um moinho de cimento, fora dos parâmetros operacionais ideais, pode ocasionar desequilíbrio no sistema, o que, por sua vez, resulta em elevação anômala nas vibrações do corpo do equipamento. Do ponto de vista operacional, esse desequilíbrio impõe a necessidade de redução na alimentação do moinho, o que acarreta maior consumo de energia e queda na eficiência produtiva. Adicionalmente, tal condição também favorece o desgaste de componentes estruturais, como as paletas do classificador e as capas dos rolos de moagem.

Com o objetivo de compreender as dinâmicas relacionadas ao aumento do diferencial de pressão em um moinho de cimento, o presente estudo busca compreender as variáveis associadas a esse fenômeno. Para isso, foi conduzida uma análise de correlações, avaliando o grau de relação entre as variáveis por meio das correlações de Pearson e de Spearman. Adicionalmente, foi implementada uma modelagem preditiva baseada em aprendizado de máquina, utilizando o algoritmo *Random Forest Regressor*.

## 4 TRABALHOS RELACIONADOS

Este capítulo tem como objetivo apresentar uma revisão da literatura sobre os estudos e pesquisas que abordam temas semelhantes ao presente trabalho.

### 4.1 Estudo de Variáveis que Afetam a Vibração do Corpo de um Moinho Vertical de Rolos (MENESES, 2023)

A monografia em questão analisa as variáveis que influenciam a vibração de moinhos verticais de rolos na produção de cimento Portland. A pesquisa investiga uma série de variáveis operacionais, como dosagens de matérias-primas, pressão de moagem, injeção de água, rotação do classificador, potência do acionamento da mesa, temperatura de entrada e saída do moinho, arraste e recirculação dos gases, diferencial de pressão do moinho e diferencial de pressão do filtro principal.

Este estudo foi fundamental para a abordagem deste trabalho, pois proporcionou uma melhor compreensão dos fenômenos dinâmicos em moinhos verticais e serviu como base para identificar as variáveis que impactam o aumento do diferencial de pressão. A metodologia adotada, que envolveu a utilização de métodos de análise de dados como defasagem temporal (lag) e média móvel, foi similar à empregada neste trabalho, sendo aplicada para suavização e compreensão do comportamento das variáveis. Adicionalmente, realizaram-se estudos de correlação e predição por meio de um modelo de regressão, que serviram de referência para a construção da metodologia utilizada neste estudo.

O presente trabalho tem como objetivo analisar as variáveis que impactam a vibração do corpo de um moinho vertical, buscando desenvolver estratégias para minimizar os efeitos de uma operação instável, como o aumento do consumo de energia e o desgaste excessivo de componentes, particularmente a solda da mesa e das capas dos rolos. Para tanto, foram aplicadas técnicas de análise de dados, como regressão linear múltipla, lag e média móvel, além de realizar uma análise exploratória da base de dados coletada, com foco na produção de cimento CP II F-32.

Apesar das interferências dos operadores, que geraram correlações espúrias, o estudo contribui para um melhor entendimento das variáveis operacionais do moinho. Embora o modelo de predição utilizado apresente algumas limitações, as conclusões indicam possibilidades de melhorias na modelagem, incluindo a adoção de técnicas de aprendizado profundo, que seriam

capazes de lidar com as relações não lineares e as dependências entre as variáveis, o que poderá aprimorar a precisão dos resultados.

#### **4.2 *Power-draw prediction by random forest based on operating parameters for an industrial ball mill (BELMAJDOUB, 2019)***

O artigo investiga a estimativa do consumo energético de um moinho de bolas industrial por meio de um modelo estatístico baseado em *Random Forest*. O estudo tem como objetivo desenvolver um modelo preditivo a partir de parâmetros operacionais, aprimorando a precisão da previsão do consumo de energia.

A pesquisa utiliza uma base de dados de um moinho de bolas em uma planta de beneficiamento de minério de ferro. Foram aplicados modelos de *Random Forest* e regressão tradicional para identificar padrões e relações entre o consumo energético e variáveis operacionais, como vazão de alimentação, velocidade do moinho, porcentagem de sólidos na polpa, carga de bolas e taxa de enchimento. A comparação dos métodos revelou que o *Random Forest* apresentou um coeficiente de determinação significativamente superior ( $R^2 = 0,98$ ) em relação aos modelos convencionais, que atingiram valores próximos a 0,60.

Além da melhoria na precisão preditiva, o estudo analisou a influência de cada parâmetro operacional no consumo de energia, permitindo uma compreensão mais detalhada dos fatores envolvidos. Os resultados apontam que o uso de modelos estatísticos avançados pode contribuir para a otimização do desempenho energético de moinhos industriais, fornecendo subsídios para ajustes operacionais mais eficientes.

#### **4.3 *Efficient machine learning model to predict fineness, in a vertical raw meal of Morocco cement plant (TOHRY S. CHEHREH CHELGANI, 2022)***

O artigo apresenta um modelo de aprendizado de máquina voltado à previsão da finura da mistura crua em um moinho vertical em uma planta de cimento localizada no Marrocos. O principal objetivo do estudo é desenvolver um sensor virtual capaz de estimar a finura do material, utilizando dados operacionais reais da planta.

Para o tratamento dos dados, os autores adotaram uma abordagem robusta, inicialmente empregando métodos de identificação de outliers multivariados, os quais foram posteriormente validados por especialistas do processo. A pesquisa selecionou oito parâmetros

de processo e quatro de qualidade do material como variáveis significativas para a construção do modelo preditivo.

Dentre as técnicas de aprendizado de máquina avaliadas, os autores utilizaram redes neurais artificiais com regularização bayesiana, algoritmos de Levenberg-Marquardt e regressão por vetor de suporte (SVR). A regularização bayesiana foi aplicada com o objetivo de evitar o sobreajuste e aprimorar a capacidade de generalização do modelo. O algoritmo de Levenberg-Marquardt foi empregado para ajustar os parâmetros das redes neurais artificiais, visando otimizar o desempenho do modelo.

Adicionalmente, o desempenho do modelo baseado em redes neurais foi comparado com o de um modelo de regressão por vetor de suporte, com o intuito de identificar a abordagem mais eficaz. O modelo de redes neurais com regularização bayesiana obteve o melhor desempenho, destacando-se como uma solução eficiente para a previsão da finura do material em condições industriais.

## 5 METODOLOGIA

Este capítulo apresenta os procedimentos e técnicas empregados na condução do presente estudo, detalhando as etapas selecionadas para analisar a influência das variáveis sobre o diferencial de pressão em um moinho vertical de rolos. A Figura 4 ilustra um fluxograma que descreve a metodologia adotada nesta pesquisa, abrangendo desde a coleta de dados até a etapa de predição.

Figura 4 – Fluxograma da metodologia do estudo.



Fonte: Elaborado pela autora (2024).

### 5.1 Coleta de dados

Elaborou-se um banco de dados contendo informações referentes ao período de operação do moinho, compreendido entre outubro de 2024 e janeiro de 2025, com registros coletados a cada 30 segundos, totalizando 300.240 informações, extraídos do software de gerenciamento PI System. A seleção dos dados foi guiada pelas variáveis consideradas pertinentes ao tema, conforme expostas na fundamentação teórica deste estudo, e definidas com base na experiência dos operadores da planta.

### 5.2 Pré-processamento de dados

Para realizar a análise proposta neste estudo, foi fundamental realizar uma análise exploratória dos dados, a fim de preparar adequadamente a base de dados. Esse processo envolve a aplicação de métodos estatísticos para extrair informações relevantes sobre o processo produtivo. Os métodos utilizados foram a suavização dos dados por meio da média móvel e o ajuste de defasagem (lag).

#### 5.2.1 Crítica aos dados

A crítica aos dados consiste em um processo de avaliação e análise rigorosa dos dados coletados, com o objetivo de identificar a qualidade, erros e inconsistências. Nesta etapa,

iniciou-se a eliminação de parâmetros irrelevantes e *outliers* do repositório, incluindo valores nulos que indicavam paradas do equipamento para manutenção preventiva, valores duplicados e variáveis ausentes.

### 5.2.2 Média móvel

A suavização de dados é um processo utilizado para reduzir o impacto de flutuações aleatórias ou ruídos em um conjunto de dados, com a intenção de identificar tendência ou padrões mais consistentes e relevantes. Tal processo é comumente utilizado na análise de séries temporais, as quais os dados podem apresentar variações irregulares e eventos pontuais que não refletem o comportamento geral.

O modelo de média móvel é uma técnica amplamente utilizada para suavizar séries temporais para um processo de previsão. O valor previsto é definido como uma combinação simples de média de valores reais recentes pertencente a uma janela temporal. Dentre as várias abordagens disponíveis para seu cálculo, este estudo adotou a média móvel simples.

Para a previsão com uma janela de três valores, calcula-se a soma dos três dados e divide-se por três para obter a média. A cada nova observação, o valor mais antigo é descartado e o novo valor é incorporado, atualizando assim o conjunto de dados utilizado para o cálculo da média móvel. Esse processo contínuo permite ajustar a previsão de acordo com os valores mais recentes, mantendo a análise centrada nas últimas observações.

$$MMS_{t+1} = \frac{X_t + X_{t-1} + X_{t-2} + \dots + X_{t-n+1}}{n} \quad (5.1)$$

A Equação 5.1 representa uma fórmula geral para média móvel simples, onde  $MMS_{t+1}$  é o valor da média móvel para período de tempo  $t + 1$ ,  $n$  é o número de períodos da janela e a sequência de  $X_t + X_{t-1} + X_{t-2} + \dots + X_{t-n+1}$  são os valores de períodos passados. Este procedimento foi implementado para cada observação da variável explicativa, resultando na construção de um conjunto de dados suavizados, o qual foi posteriormente integrado ao modelo de previsão da variável resposta.

### 5.2.3 Lag

O termo Lag ou desfasamentos descreve o intervalo temporal entre eventos em uma série temporal, refletindo o número de períodos durante os quais uma variável explicativa

impacta a variável resposta. No contexto deste estudo, um Lag corresponde a um intervalo de 30 segundos. Assim, um desfasamento de 20 períodos implica um atraso temporal de 10 minutos. Este impacto é avaliado a partir da análise da variação da variável resposta em função das modificações na variável independente.

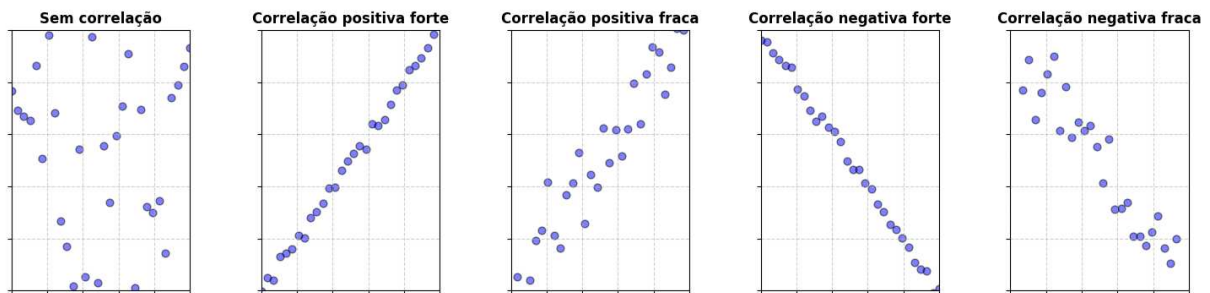
### 5.3 Correlações

As correlações indicam o grau de relacionamento entre variáveis, descrevendo sua associação e variação conjunta. Esses coeficientes expressam, por meio de um único valor, o grau da dependência entre as variáveis analisadas. Geralmente, os coeficientes variam entre 0 e 1, ou entre -1 e +1, sendo que valores próximos de zero indicam ausência de associação entre as variáveis (BUSSAB; MORETTIN, 2010). Neste estudo, foram considerados os coeficientes de correlação de Pearson, Spearman e Informação Mútua.

#### 5.3.1 Correlação de Pearson

O coeficiente de correlação de Pearson mede a direção e a intensidade da relação linear entre duas variáveis quantitativas, assumindo que ambas seguem uma distribuição normal (MOORE; MCCABE, 2009). Nesse modelo, variações na variável independente geram impactos proporcionais na variável dependente, permitindo representar a relação por uma equação linear de primeiro grau, cuja inclinação indica uma associação positiva ou negativa. A Figura 5 exemplifica essa afirmativa.

Figura 5 – Exemplos de gráficos de dispersão de correlação de Pearson



Fonte: Elaborado pela autora (2024)

De acordo com Johnson e Wichern (2007), a correlação de Pearson é formalmente definida como a razão entre a covariância de duas variáveis e o produto de seus respectivos desvios padrões. Essa relação é descrita pela Equação 5.2.

$$r_{X,Y} = \frac{\text{cov}(X,Y)}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (5.2)$$

Onde X e Y representam dois conjuntos de variáveis distintas,  $\text{cov}(X, Y)$  corresponde à covariância entre elas e  $n$  a quantidade de pares de valores denotados por  $(x_1, y_1), \dots, (x_n, y_n)$ .

A covariância, também denominada variação conjunta, é uma medida que quantifica o grau de interdependência quantitativa entre duas variáveis aleatórias. Seu cálculo baseia-se na média dos produtos das diferenças entre cada valor individual e a média das respectivas variáveis, conforme definido pela Equação 5.3.

$$\text{cov}(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n} \quad (5.3)$$

Segundo Cover e Thomas (2006), a interpretação dos coeficientes de correlação é realizada da seguinte forma: valores absolutos entre 0 e 0,3 (ou 0 e -0,3) indicam correlações consideradas biologicamente desprezíveis; entre 0,31 e 0,5 (ou -0,31 e -0,5) representam correlações fracas; entre 0,51 e 0,7 (ou -0,51 e -0,7) correspondem a correlações moderadas; entre 0,71 e 0,9 (ou -0,71 e -0,9) indicam correlações fortes; e valores superiores a 0,9 (ou inferiores a -0,9) são classificados como correlações muito fortes. A Tabela 1 apresenta essa classificação de forma consolidada.

Tabela 1 – Interpretação do coeficiente de correlação de Pearson

Intervalo de Coeficiente de Correlação	Classificação
0 a 0,1 (ou 0 a -0,1)	Desprezível
-0,1 a 0,3 (ou 0 a -0,3)	Muito fraca
0,31 a 0,5 (ou -0,31 a -0,5)	Fraca
0,51 a 0,7 (ou -0,51 a -0,7)	Moderada
0,71 a 0,9 (ou -0,71 a -0,9)	Forte
0,9 a 1 (-0,9 a -1)	Muito forte
= 1 (ou = -1)	Correlação linear perfeita

Fonte: Adaptado de Cover e Thomas (2006)

### 5.3.2 Correlação de informação mútua

A correlação de informação mútua ou apenas correlação mútua é utilizada para mensurar a dependência ou interdependência entre duas variáveis ou conjuntos de dados. Em contraste com a correlação de Pearson, a informação mútua permite a interpretação de relações

complexas, incluindo aquelas que não apresentam caráter linear. A Equação 5.4 apresenta a metodologia para o cálculo dessa correlação, conforme descrito por Cover e Thomas (2006).

$$I(X;Y) = \sum_{x,y} p(x,y) \log \left( \frac{p(x,y)}{p(x)p(y)} \right) \quad (5.4)$$

A probabilidade conjunta, denotada por  $p(x,y)$  é uma medida que quantifica a probabilidade de ocorrência simultânea de duas variáveis aleatórias, refletindo a dependência entre elas. Já as probabilidades marginais, representadas por  $p(x)$  e  $p(y)$ , correspondem às distribuições individuais de  $x$  e  $y$ , respectivamente, e indicam a probabilidade de ocorrência de cada variável de forma independente, sem considerar as interações com outras variáveis

A correlação mútua não assume valores negativos. A dependência mútua é considerada quando os valores são positivos, sendo que, à medida que se aproximam do infinito, indicam que uma variável pode ser perfeitamente predita pela outra. A independência total é caracterizada quando o valor da correlação mútua é igual a zero.

### 5.3.3 Correlação de Spearman

A correlação de Spearman, representado pela letra grega  $\rho$  (rô), é utilizada para avaliar a associação monotônica entre duas variáveis quantitativas que não necessariamente seguem uma distribuição normal ou possuem categorias ordenadas (SPEARMAN, 1904). Esse coeficiente permite identificar relações não lineares que, ainda assim, mantêm uma ordem sistemática, ou seja, quando uma variável maior tende a estar associada a uma variável também maior, ou, alternativamente, a uma variável menor.

O coeficiente em questão é calculado de maneira semelhante ao coeficiente de correlação de Pearson. A principal distinção reside no fato de que, nesse caso, a covariância é obtida com base nos postos ordenados das variáveis, em vez de seus valores originais. A Equação 5.5 formaliza o cálculo desse coeficiente.

$$\rho_{rg_X, rg_Y} = \frac{cov(rg_X, rg_Y)}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (5.5)$$

Onde  $cov(rg_X, rg_Y)$  representa a covariância entre os postos ordenados dos conjuntos analisados. A interpretação dos valores obtidos segue o mesmo princípio aplicado ao coeficiente de correlação de Pearson,

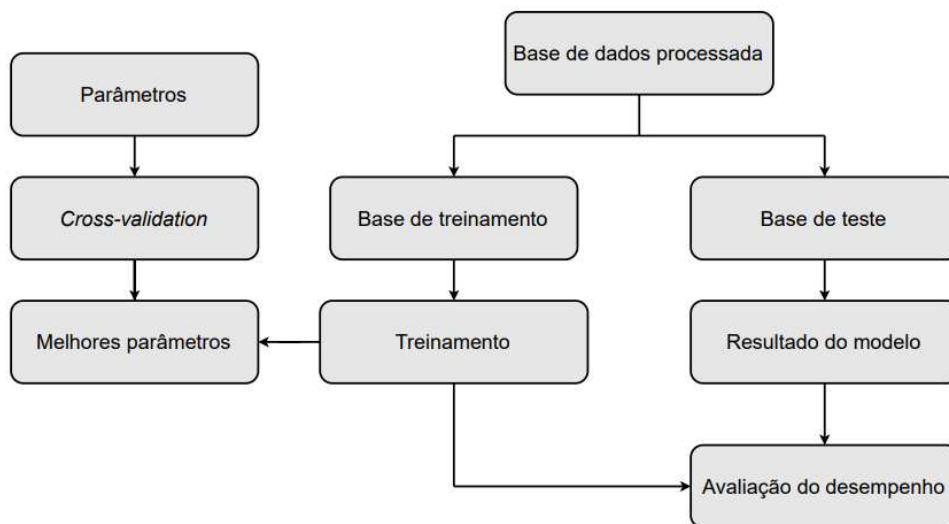
## 5.4 Aprendizado de máquina

O *machine learning*, ou aprendizado de máquina, é uma subárea da inteligência artificial que capacita os computadores a aprenderem a partir de dados e experiências anteriores, sem a necessidade de programação explícita para tal (GÉRON, 2019). Os modelos de aprendizado de máquina identificam padrões nos dados e utilizam esses padrões para realizar previsões ou tomar decisões. Para uma melhor compreensão, considera-se a afirmação de Mitchell (1997, p. 2).

Diz-se que um programa de computador aprende pela experiência E em relação a algum tipo de tarefa T e alguma medida de desempenho P se o seu desempenho em T, conforme medido por P, melhora com a experiência E.

A modelagem refere-se ao processo de desenvolvimento de um modelo de *machine learning* a partir de um conjunto de dados históricos, em que se seleciona um modelo adequado que dará origem a um algoritmo, ou a um conjunto de algoritmos, a ser empregado durante o treinamento. Essa etapa tem como objetivo capacitar o modelo a identificar padrões nos dados, resultando em um modelo treinado que assimila o conhecimento extraído das informações históricas. A Figura 6 demonstra as etapas seguidas no processo de modelagem.

Figura 6 – Fluxograma das etapas de modelagem.



Fonte: Elaborado pela autora (2024).

Existem dois tipos principais de aprendizado de máquina: o supervisionado, que utiliza dados rotulados para treinar modelos, e o não supervisionado, que identifica padrões em dados não rotulados. No aprendizado supervisionado, as árvores de decisão segmentam os dados

em subconjuntos com base em condições específicas, nas quais os nós representam perguntas ou testes, e as folhas correspondem às previsões ou decisões resultantes.

O *Random Forest Regressor* configura-se como uma técnica de aprendizado supervisionado aplicada à resolução de problemas de previsão de valores contínuos (BREIMAN, 1996). Nesse contexto, cada árvore de decisão presente na floresta realiza uma previsão numérica, e a predição final é determinada pela média das saídas de todas as árvores. O principal objetivo deste método é reduzir a variabilidade e aprimorar a precisão das previsões.

$$\hat{y} = \frac{1}{N} \sum_{i=1}^N h_i(x) \quad (5.6)$$

A Equação 5.7 apresentada descreve o cálculo da predição final para regressão, no qual  $N$  é o número total de árvores no modelo e  $h_i(x)$  é a predição feita pela  $i$ -ésima árvore para a amostra de teste. Para cada nó da árvore, o algoritmo seleciona uma variável dos dados de entrada para dividir os dados em subconjuntos.

A parametrização de um modelo impacta seu comportamento durante o treinamento e a geração de previsões. A definição de parâmetros em um modelo possibilita o equilíbrio da variância, evitando o *overfitting* (sobreajuste) ou *underfitting* (subajuste). O sobreajuste ocorre quando o modelo se ajusta excessivamente aos dados de treinamento, memorizando padrões específicos e ruídos, enquanto o subajuste acontece quando o modelo é muito simples para capturar padrões relevantes, resultando em baixo desempenho.

Os principais parâmetros incluem o número de árvores, a profundidade máxima das árvores, o número mínimo de amostras para realizar uma divisão em um nó e o número mínimo de amostras exigidas em um nó folha (GÉRON, 2019). Esses parâmetros podem ser ajustados para otimizar o desempenho do modelo, conforme as características dos dados e do problema em questão.

A escolha dos parâmetros e a definição das métricas de avaliação são fundamentais para a análise do desempenho do modelo. Conforme indicado por Bishop (2006), a validação cruzada é uma técnica crucial na avaliação de modelos, pois permite verificar a capacidade de generalização dos mesmos. A análise de desempenho do modelo em questão deve ser conduzida por meio de métricas de regressão, como o erro absoluto médio escalonado (MASE), a raiz do erro quadrático médio (RMSE) e o erro absoluto médio (MAE).

As Equações 5.7, 5.8 e 5.9 correspondentes a essas métricas são as seguintes:

$$MASE = \frac{\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|}{\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|} \quad (5.7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5.8)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5.9)$$

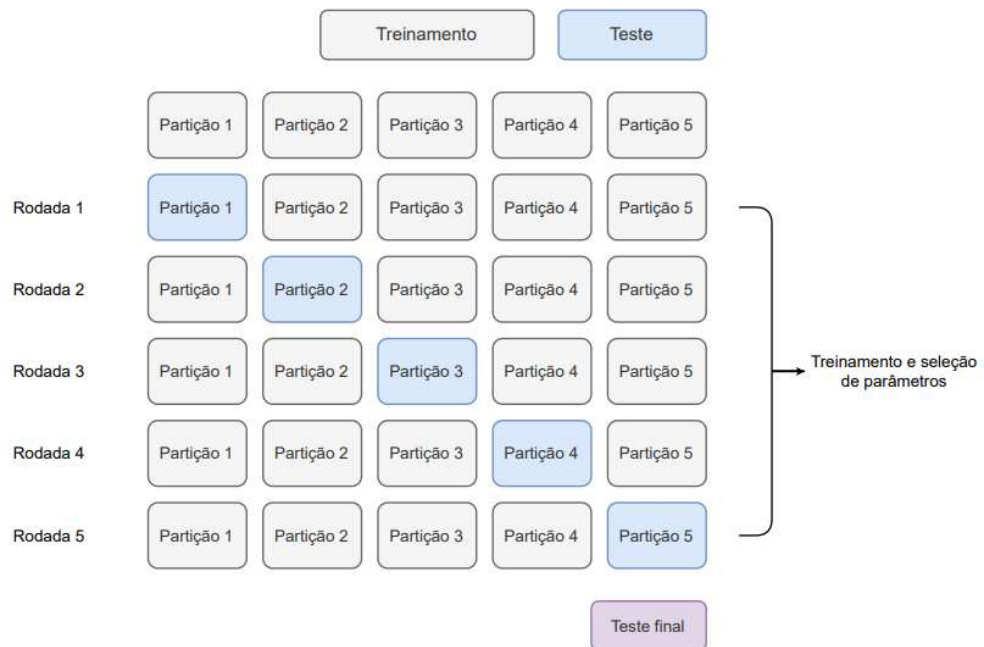
Onde  $y_i$  representa o valor real,  $\hat{y}_i$  o valor previsto e  $n$  o número total de observações.

Essas métricas devem ser determinadas com base em testes realizados em um conjunto de dados distinto daquele utilizado no treinamento. Dessa forma, o modelo será avaliado com duas bases de dados: uma para o treinamento e outra para o teste, garantindo que os dados da base de teste sejam completamente independentes dos dados empregados no treinamento. O objetivo do teste é analisar o desempenho do modelo com dados inéditos, simulando uma situação de aplicação prática.

A prática mencionada no parágrafo anterior é denominada validação cruzada ou *K-fold cross validation*. Esse procedimento consiste na divisão do conjunto de dados em  $K$  grupos, denominados *folds*, realizando uma iteração para cada grupo. Em cada iteração, um grupo é utilizado para testes, enquanto os demais são empregados no treinamento do modelo. . A Figura 7 ilustra o processo da validação cruzada para o caso em que  $K=5$ .

Normalmente, aproximadamente 20% dos dados são reservados para o teste. O modelo é treinado com os dados de treinamento, avaliado com os dados de teste, a métrica é registrada e o modelo é descartado. Após cada iteração, o próximo grupo é selecionado. Ao final do processo, as métricas de todas as iterações são analisadas.

Figura 7 – Ilustração de cada tipo de partição para cross-validation.



Fonte: Adaptado de Scikit-learn (2024).

## 6 RESULTADOS

Este capítulo apresenta os resultados da análise dos dados coletados sobre o diferencial de pressão na produção de cimento em um moinho de rolos vertical, conforme detalhado na metodologia deste estudo. A seção está estruturada em subseções, que abordam a análise exploratória dos dados, a análise da defasagem temporal, as correlações de Pearson, Informação Mútua e de Spearman, e modelagem com *Random Forest Regressor*.

### 6.1 Análise exploratória de dados

A análise exploratória de dados consiste na avaliação inicial de um conjunto de dados, com o objetivo de identificar padrões e relações relevantes antes da aplicação de métodos de análise mais avançados ou modelos estatísticos. Nesta etapa, foram elaborados gráficos de séries temporais, histogramas e boxplots para examinar o comportamento do diferencial de pressão e das variáveis associadas, além de identificar a presença de *outliers*.

A Tabela 2 apresenta as medidas estatísticas das variáveis dentro do período analisado, incluindo a média, o desvio padrão e variância. Este último utilizado para quantificar a dispersão dos dados em relação à média. Analisando a tabela, verificou-se que os maiores valores de desvio padrão ocorreram para a velocidade do classificador, temperatura de entrada e alimentação total. A variância também se mostrou expressiva nestas variáveis, acompanhadas pelo damper 02.

Tabela 2 – Medidas estatísticas sobre as variáveis analisadas

Variável	Unidade	Média	Desvio padrão	Variância
Diferencial de pressão	mbar	25,04	2,19	4,81
Pressão de entrada	mbar	6,08	2,41	5,83
Pressão de saída	mbar	33,64	2,07	4,3014
Temperatura de entrada	°C	123,32	16,62	276,27
Temperatura de saída	°C	103,27	3,54	12,56
Velocidade do classificador	rmp	639,05	48,67	2368,87
Injeção de água	m <sup>3</sup> /h	1,07	0,95	0,90
Pressão de moagem	bar	113,14	6,51	42,43
Alimentação total	t/h	70,88	13,24	175,389
Damper 02	%	44,52	6,58	43,37

Fonte: Elaborado pela autora (2024)

Esse fato indica uma maior variabilidade nos dados dessas variáveis, resultante das alterações nas condições operacionais do moinho. Isso ocorre porque existem limites operacionais ideais para a produção de cimento, os quais dependem da dosagem da matéria-

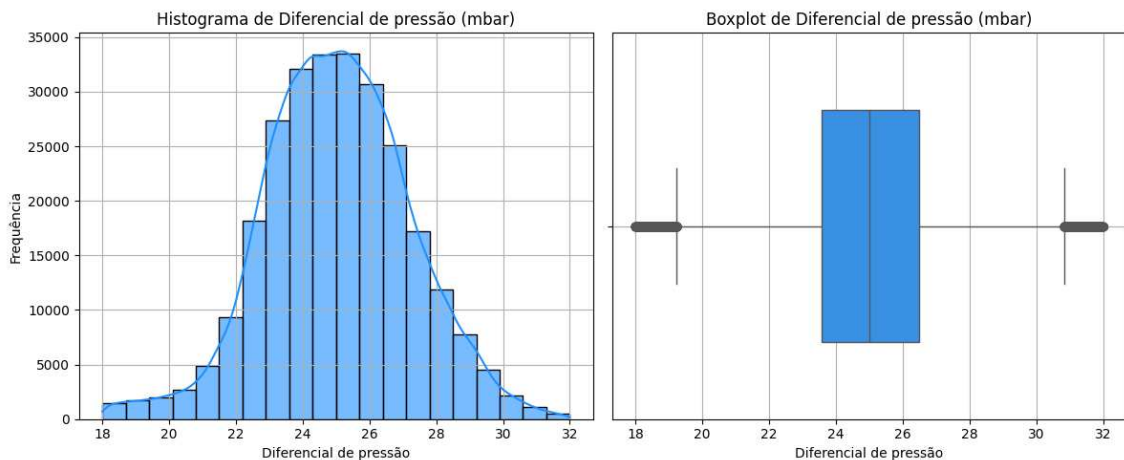
prima. Com base nisso, constatou-se a necessidade de aplicar filtragens nas variáveis estudadas, com o objetivo de eliminar ruídos que poderiam comprometer os estudos de correlação e predição.

A análise da interceptação dos dados foi realizada por meio de gráficos apresentados no início desta seção, os quais indicaram as regiões com maior densidade de informações, bem como a presença de valores atípicos com uso de histogramas e bloxpot. Para garantir maior clareza, esta subseção foi organizada em tópicos para cada variável.

### 6.1.1 *Análise do diferencial de pressão*

A Figura 8 o histograma e o bloxpot relacionados à variável resposta, constatou-se uma maior frequência no intervalo de 18 a 32 *mbar*, expresso pelos valores dos limites inferior e superior do bloxpot. O mesmo processo de análise da variável resposta foi aplicado as variáveis explicativas, levando em consideração os limites inferiores e superiores do bloxpot e distribuição no histograma.

Figura 8 – Gráficos histograma e bloxpot do diferencial de pressão.

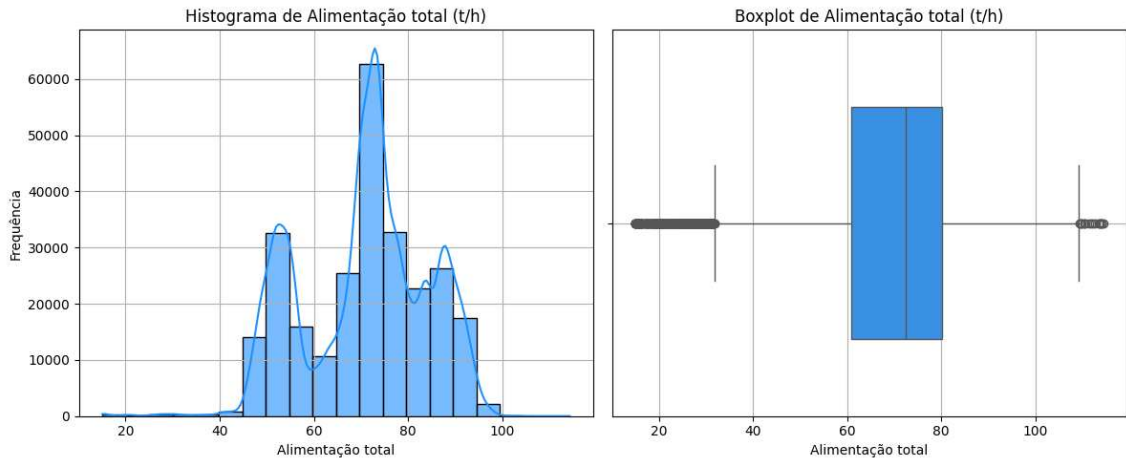


Fonte: Elaborado pela autora (2024).

### 6.1.2 *Análise da alimentação total*

A alimentação do moinho influencia diretamente a pressão de saída, pois define a quantidade de material que deixa o equipamento. Esse parâmetro é dosado para evitar vibrações no corpo do moinho e garantir a qualidade do produto final. Assim, a definição de um intervalo adequado é essencial para a estabilidade da análise. Conforme ilustrado na Figura 9, a filtragem dos dados de alimentação total considera o intervalo de 40 a 115 t/h, correspondendo aos limites dos boxplots.

Figura 9 – Gráficos histograma e bloxpot da alimentação total.

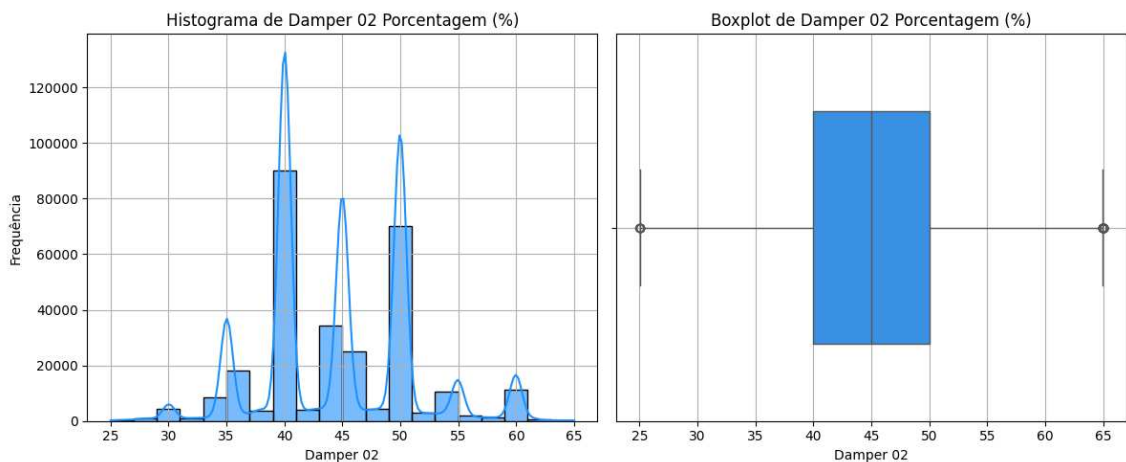


Fonte: Elaborado pela autora (2024).

### 6.1.3 Análise do arraste de gases

A Figura 10 apresenta uma variação expressiva dos dados, variando entre 25% e 65%, conforme os limites dos boxplots.

Figura 10 – Gráficos histograma e bloxpot do *damper 02*.

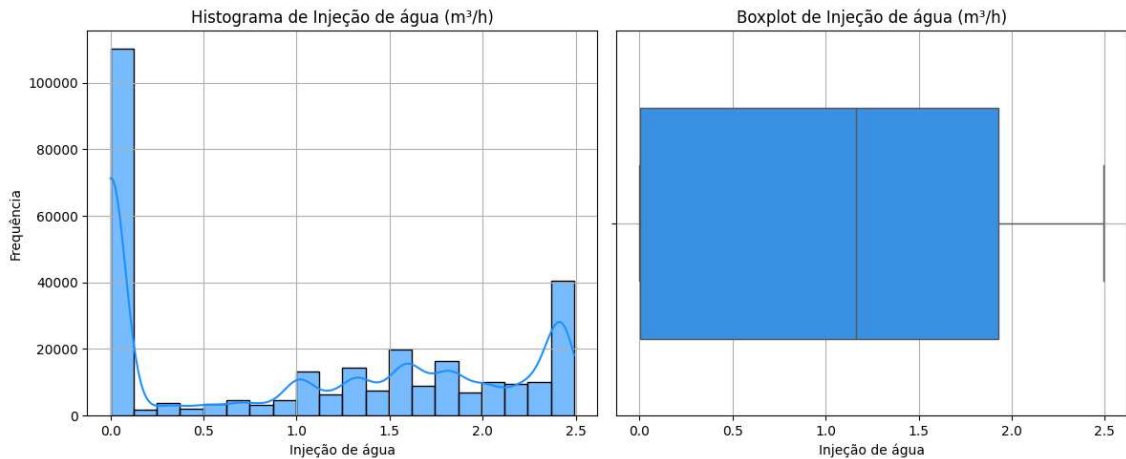


Fonte: Elaborado pela autora (2024).

### 6.1.4 Análise da injeção de água

A Figura 11 refere-se à variação da pressão de entrada do moinho, na qual observa-se uma distribuição homogênea entre 1 e 2,5 m<sup>3</sup>/h, sendo esse intervalo considerado adequado para a filtragem da variável.

Figura 11 – Gráficos histograma e bloxpot da injeção de água.

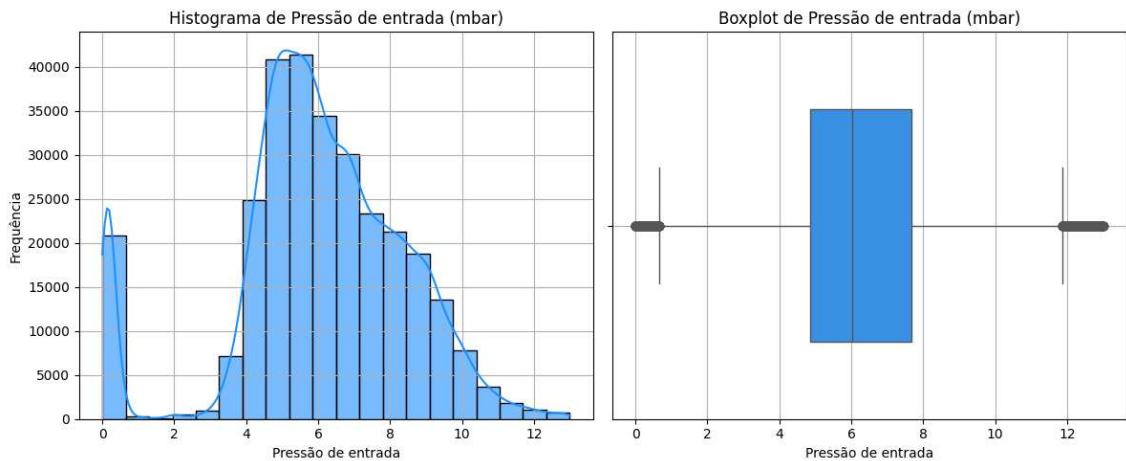


Fonte: Elaborado pela autora (2024).

### 6.1.5 Análise da pressão de entrada

A Figura 12 ilustra a variação da pressão de entrada do moinho, na qual se verifica que o primeiro quartil corresponde a 4,86 *mbar* e o terceiro quartil a 7,66 *mbar*. Os limites inferior e superior da distribuição são, respectivamente, 11,86 *mbar* e 0,66 *mbar*, indicando que os *outliers* estão fora desse intervalo. Diante disso, decidiu-se aplicar uma intercepção para valores entre 1 e 12 *mbar*.

Figura 12 – Gráficos histograma e bloxpot da pressão de entrada.



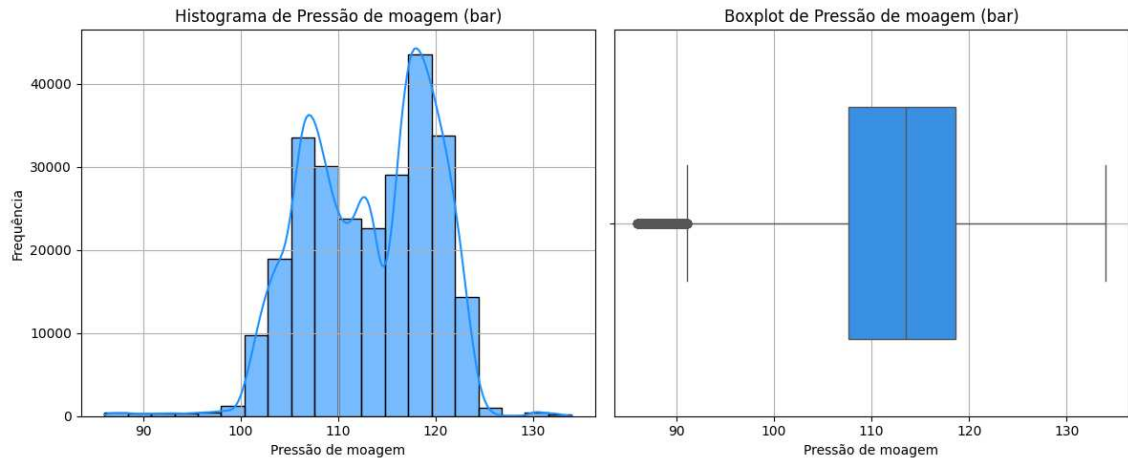
Fonte: Elaborado pela autora (2024).

### 6.1.6 Análise da pressão de moagem

A Figura 13 demonstra o histograma e o boxplot da pressão de moagem. O primeiro quartil corresponde a 105,47 bar, enquanto o terceiro quartil é de 118,07 bar. O limite inferior é

de 86.57 bar, e o limite superior, de 136.97 bar. Diante disso, escolheu-se interceptar os valores entre os 86 a 137 bar.

Figura 13 – Gráficos histograma e bloxpot da pressão de moagem.

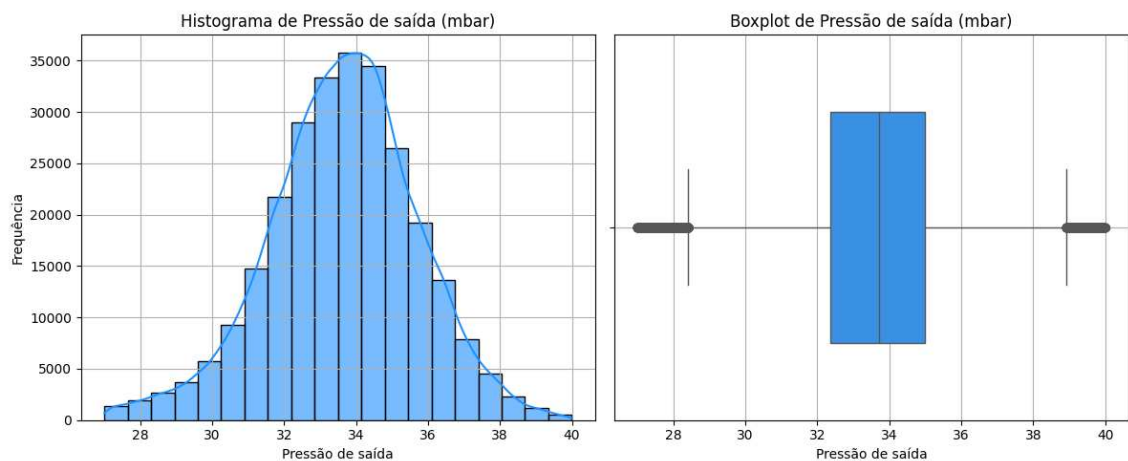


Fonte: Elaborado pela autora (2024).

### 6.1.7 Análise da pressão de saída

A Figura 14 ilustra o histograma e o boxplot da variável pressão de saída. O primeiro quartil apresenta valor de 31,66 *mbar*, enquanto o terceiro quartil é de 34,83 *mbar*. Os limites inferior e superior são, respectivamente, 26,9 e 39,6 *mbar*. Assim, optou-se por restringir os dados ao intervalo entre 27 e 40 *mbar*.

Figura 14 – Gráficos histograma e bloxpot da temperatura de entrada.

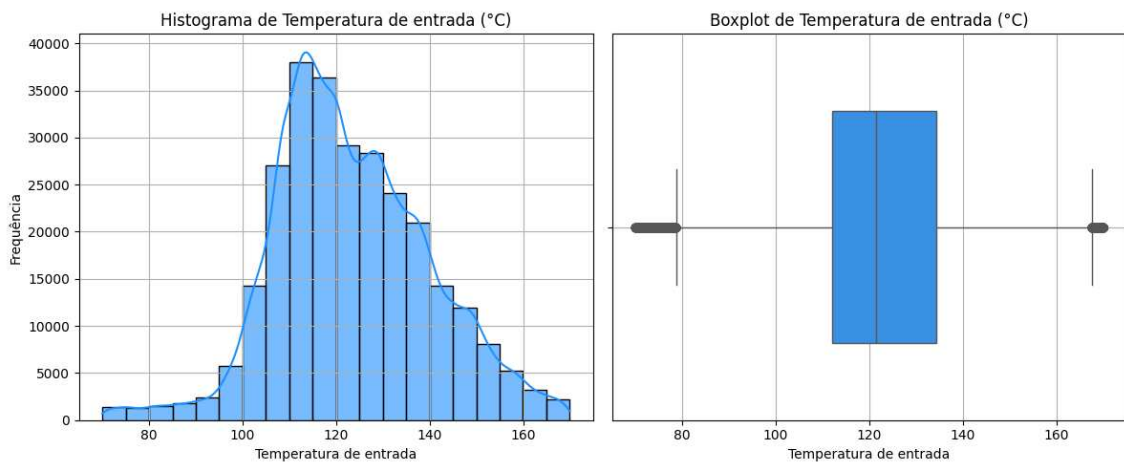


Fonte: Elaborado pela autora (2024).

### 6.1.8 Análise da temperatura de entrada

A Figura 15 exibe o histograma e o boxplot desta variável. Inicialmente, observa-se uma maior concentração de dados na faixa de 100 a 160°C. O primeiro quartil encontra-se em 167,60°C, e o terceiro quartil, em 133,86°C. Os limites inferior e superior são, respectivamente, 78,63°C e 167,60°C. Considerando isso, optou-se por filtrar os dados para o intervalo de 70 a 170°C.

Figura 15 – Gráficos histograma e bloxpot da temperatura de entrada.



Fonte: Elaborado pela autora (2024).

### 6.1.9 Análise da temperatura de saída

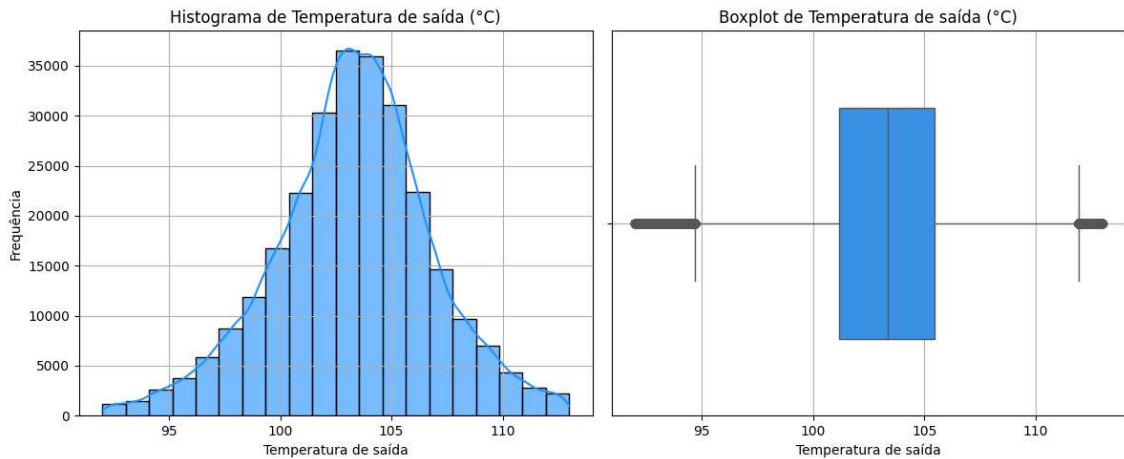
A Figura 16 mostra a variação da pressão de entrada do moinho, evidenciando que o primeiro quartil é de 100,18°C, enquanto o terceiro quartil é de 105,37°C. Os limites inferior e superior da distribuição são 92,38 e 113,17°C, respectivamente, o que sugere que os *outliers* se encontram fora desse intervalo. Com isso, optou-se por aplicar uma intercepção para valores entre 92 e 113°C.

### 6.1.10 Análise da velocidade do classificador

A Figura 17 ilustra o histograma e o boxplot desta variável da velocidade do classificador. O primeiro quartil é de 610,32 rpm, e o terceiro quartil, de 688,80 rpm. Os limites inferior e superior são, respectivamente, 492,59 rpm e 806,52 rpm. Desse modo, decidiu-se restringir os dados ao intervalo entre 500 e 800 rpm.

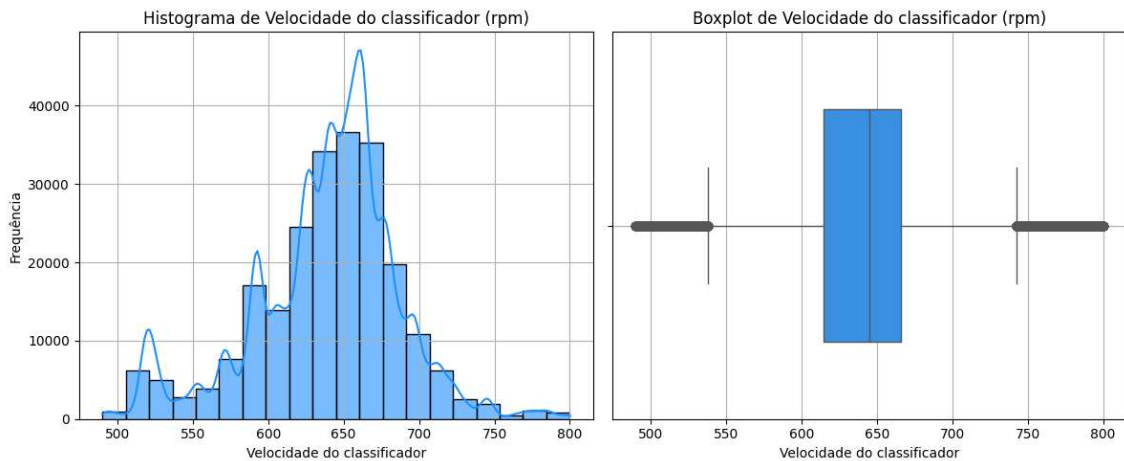
A Tabela 3 apresenta as métricas estatísticas das variáveis após a aplicação do

Figura 16 – Gráficos histograma e bloxpot da temperatura de saída.



Fonte: Elaborado pela autora (2024).

Figura 17 – Gráficos histograma e bloxpot da velocidade do classificador.



Fonte: Elaborado pela autora (2024).

processo de filtragem. Observa-se uma redução nos valores de desvio padrão e variância, o que indica maior consistência em torno da média. Embora a velocidade do classificador e a injeção de água ainda apresentem desvios elevados, optou-se por não realizar uma nova filtragem, a fim de preservar a quantidade de dados e manter o comportamento original.

## 6.2 Análise da defasagem de tempo

A análise da defasagem temporal é utilizada para compreender o atraso ou a relação entre dois conjuntos de dados ou eventos ao longo do tempo. No presente estudo, a defasagem foi definida pelo lag temporal, sendo cada unidade de lag equivalente a um intervalo de 30 segundos. Dessa forma, um atraso de 30 lags corresponde a 900 segundos, ou 15 minutos, o que representou o valor máximo considerado na análise. Para determinar a configuração mais adequada, foram avaliados os valores de lag e de média móvel que apresentassem a maior correlação, utilizando-se

Tabela 3 – Medidas estatísticas sobre as variáveis após a filtragem

Variável	Intervalo	Unidade de medida	Desvio padrão	Variância
Alimentação total	15 a 115	t/h	10,22	104,63
Damper 02	25 a 65	%	6	34,37
Diferencial de pressão	19 a 30	mbar	1,90	3,64
Injeção de água	1 a 2,5	m <sup>3</sup> /h	1	1
Pressão de entrada	1 a 12	mbar	1,77	3,14
Pressão de moagem	86 a 137	bar	6,41	41,17
Pressão de saída	27 a 40	mbar	1,86	3,47
Temperatura de entrada	70 a 170	°C	15,23	232,05
Temperatura de saída	90 a 113	°C	3,51	12,52
Velocidade do classificador	500 a 806	rpm	45,6	2087,94

Fonte: Elaborado pela autora (2024)

um intervalo de janelas de médias móveis limitado a 20.

O procedimento descrito acima foi reproduzido para as correlações de Pearson, Spearman e Informação Mútua. As Tabelas 4, 5 e 6 apresentam os valores ótimos de defasagem e média móvel para os valores máximos das correlações mencionadas. Além de descrever a classificação conforme apresentado na Tabela 1. Cumpre destacar que essas análises foram realizadas para os dados dentro dos intervalos de filtragem apresentados no tópico anterior.

A correlação de Pearson indicou os maiores valores de correlação para a maioria das variáveis explicativas, enquanto a correlação de Spearman apresentou os valores mais baixos. Além disso, a correlação mútua também apresentou valores elevados. Tal fato sugere que as relações entre as variáveis explicativas e o diferencial de pressão são predominantemente lineares, com forte dependência estatística entre elas. No entanto, a menor intensidade dos coeficientes de Spearman pode indicar a presença de padrões não lineares menos expressivos.

As variáveis relacionadas a pressão de moagem e *damper 02* apresentam os menores valores de correlação para todos os tipos de correlação. Esse comportamento pode ser justificado pela interferência de outras variáveis, uma vez que essas apresentam correlações mais fortes com outras variáveis explicativas. Uma análise mais detalhada sobre este ponto será apresentada no tópico subsequente, dedicado à discussão das correlações.

As Figuras 18 e 19 apresentam gráficos de *cross-correlation* ou correlação cruzada, que ilustram o comportamento das correlações investigadas em função do aumento do Lag para cada variável explicativa. Na maioria das variáveis, verifica-se que o incremento do Lag tende a reduzir a correlação com o diferencial de pressão, indicando uma redução na intensidade da relação linear ao longo do tempo.

A análise da correlação de Pearson apresentada nos gráficos evidencia que a influên-

Tabela 4 – Tabela de valores de lag e janela de média móvel para os valores máximos de correlação de Pearson

Variável	Corr. Pearson	Classificação da correlação	Lag	Janela
Alimentação total	0,4986	Fraca	6	20
Damper 02	-0,0828	Desprezível	1	20
Injeção de água	0,4202	Fraca	30	20
Pressão de entrada	-0,4569	Fraca	1	20
Pressão de moagem	-0,1246	Muito fraca	30	20
Pressão de saída	0,7172	Forte	1	8
Temperatura de entrada	0,2529	Muito fraca	1	20
Temperatura de saída	0,2896	Muito fraca	30	20
Velocidade do classificador	0,4280	Fraca	4	16

Fonte: Elaborado pela autora (2024)

Tabela 5 – Tabela de valores de lag e janela de média móvel para os valores máximos de correlação de Spearman

Variável	Corr. Spearman	Classificação da correlação	Lag	Janela
Alimentação total	0,4748	Fraca	6	20
Damper 02	-0,0884	Fraca	1	20
Injeção de água	0,4285	Fraca	30	20
Pressão de entrada	-0,4921	Fraca	1	20
Pressão de moagem	-0,0794	Desprezível	30	1
Pressão de saída	0,6853	Moderada	1	6
Temperatura de entrada	0,2129	Muito fraca	1	13
Temperatura de saída	0,2579	Muito fraca	1	20
Velocidade do classificador	0,4504	Fraca	3	9

Fonte: Elaborado pela autora (2024)

Tabela 6 – Tabela de valores de lag e janela de média móvel para os valores máximos da correlação de Informação Mútua

Variável	Informação Mútua	Classificação da correlação	Lag	Janela
Alimentação total	0,4403	Fraca	7	20
Damper 02	0,2130	Muito fraca	24	20
Injeção de água	0,4157	Fraca	4	20
Pressão de entrada	0,2996	Muito fraca	1	20
Pressão de moagem	0,1595	Muito fraca	7	20
Pressão de saída	0,5710	Moderada	1	20
Temperatura de entrada	0,2202	Muito fraca	2	20
Temperatura de saída	0,2023	Muito fraca	6	20
Velocidade do classificador	0,4047	Fraca	2	20

Fonte: Elaborado pela autora (2024)

cia das janelas de médias móveis é mais pronunciada nas variáveis alimentação total, pressão de saída e velocidade do classificador, nas quais se observa uma distinção clara entre as curvas

correspondentes às diferentes janelas. Ademais, a temperatura de entrada e o Damper 02 exibem uma maior dispersão nos valores de correlação, sugerindo um comportamento menos linear e mais sensível às variações decorrentes da aplicação das médias móveis.

A injeção de água e a pressão de entrada também apresentam uma variação significativa na correlação de Pearson, indicando uma influência moderada das médias móveis. Por outro lado, a pressão de moagem demonstra um comportamento mais estável, com menor variação entre as janelas analisadas.

Notadamente, a injeção de água apresenta uma variação mais acentuada na correlação de Spearman em relação à Pearson, sugerindo que essa variável pode apresentar relações monotônicas mais expressivas, embora não necessariamente lineares. O mesmo padrão é observado para a pressão de entrada, reforçando a maior sensibilidade da correlação de Spearman a padrões não lineares.

A correlação de Informação Mútua apresentou comportamento oscilatório e disperso com o aumento da defasagem temporal. Não foram identificados padrões nítidos de dispersão entre as variáveis, especialmente em relação à pressão de moagem, sugerindo uma relação fraca ou a influência de ruídos não identificados. Para a pressão de saída e a velocidade do classificador, observou-se redução gradual da Informação Mútua, indicando uma relação inicialmente forte, mas em declínio ao longo do tempo.

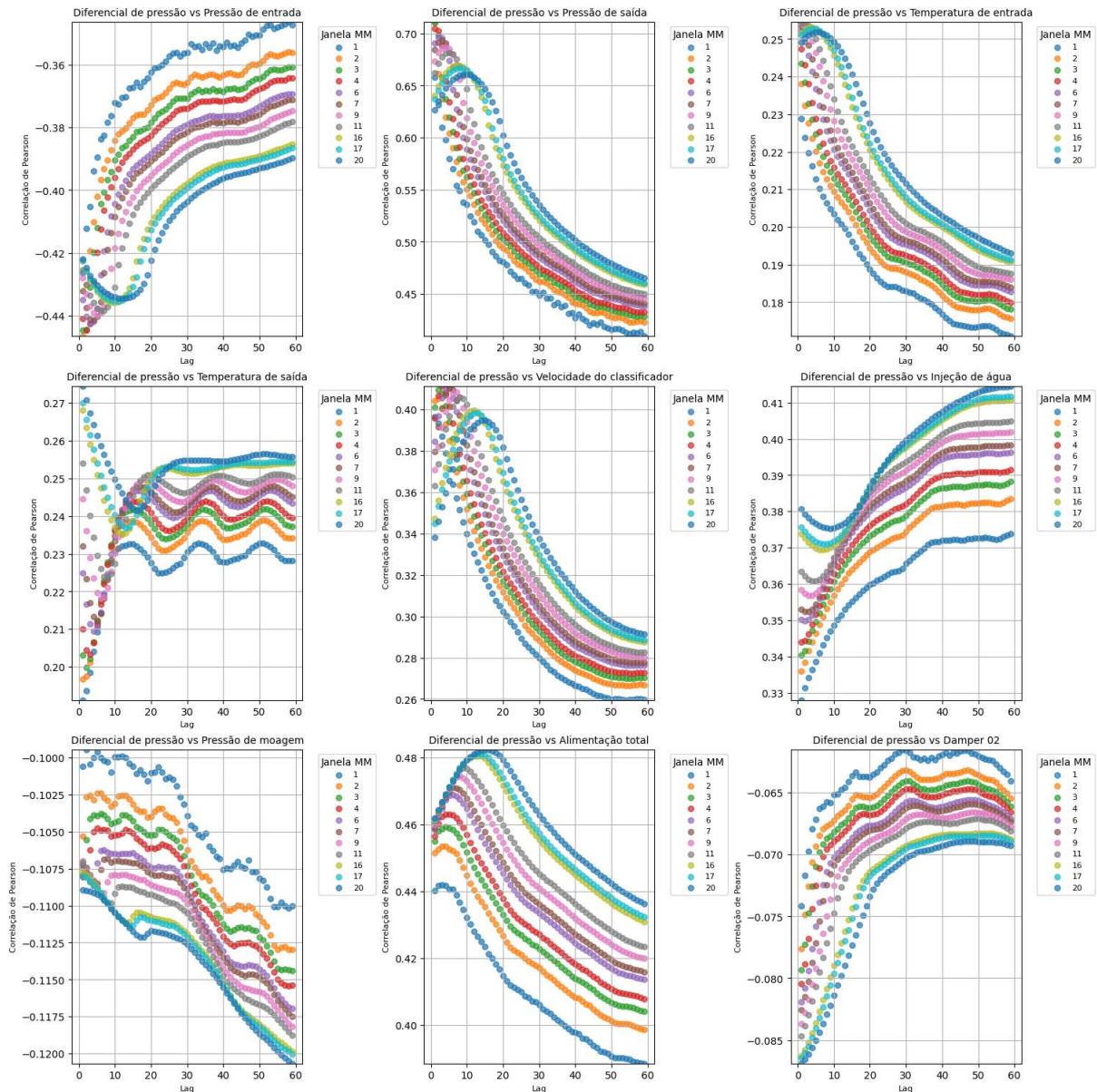
Seguindo o exposto, foram definidos os valores de lag e de janela móvel para cada variável explicativa expostos na Tabela 7. A partir dessa seleção, as etapas subsequentes do estudo se concentrarão na análise dos dados processados por meio dessas suavizações. Nos tópicos a seguir, será abordada a influência dessas suavizações nas correlações entre as variáveis e a predição será realizada com base nesses aspectos.

### **6.3 Análise das relações via correlação de Pearson**

A Figura 21 exibe os *heatmaps* ou mapas de calor correspondentes a correlação de Pearson. Gráfico (a) para os dados brutos, (b) para filtrados e (c) com aplicação das suavizações. Verifica-se um aumento nos valores absolutos das correlações com a implementação de lag e média móvel. Contudo, as variáveis pressão de moagem e *damper 02* continuam apresentando valores baixos ao final do processamento dos dados.

Esses comportamentos podem ser atribuídos ao efeito de outras variáveis sobre as mencionadas ou à natureza mais complexa dessas relações. No que se refere à pressão

Figura 18 – Correlação cruzada de Pearson para diferentes valores de média móvel

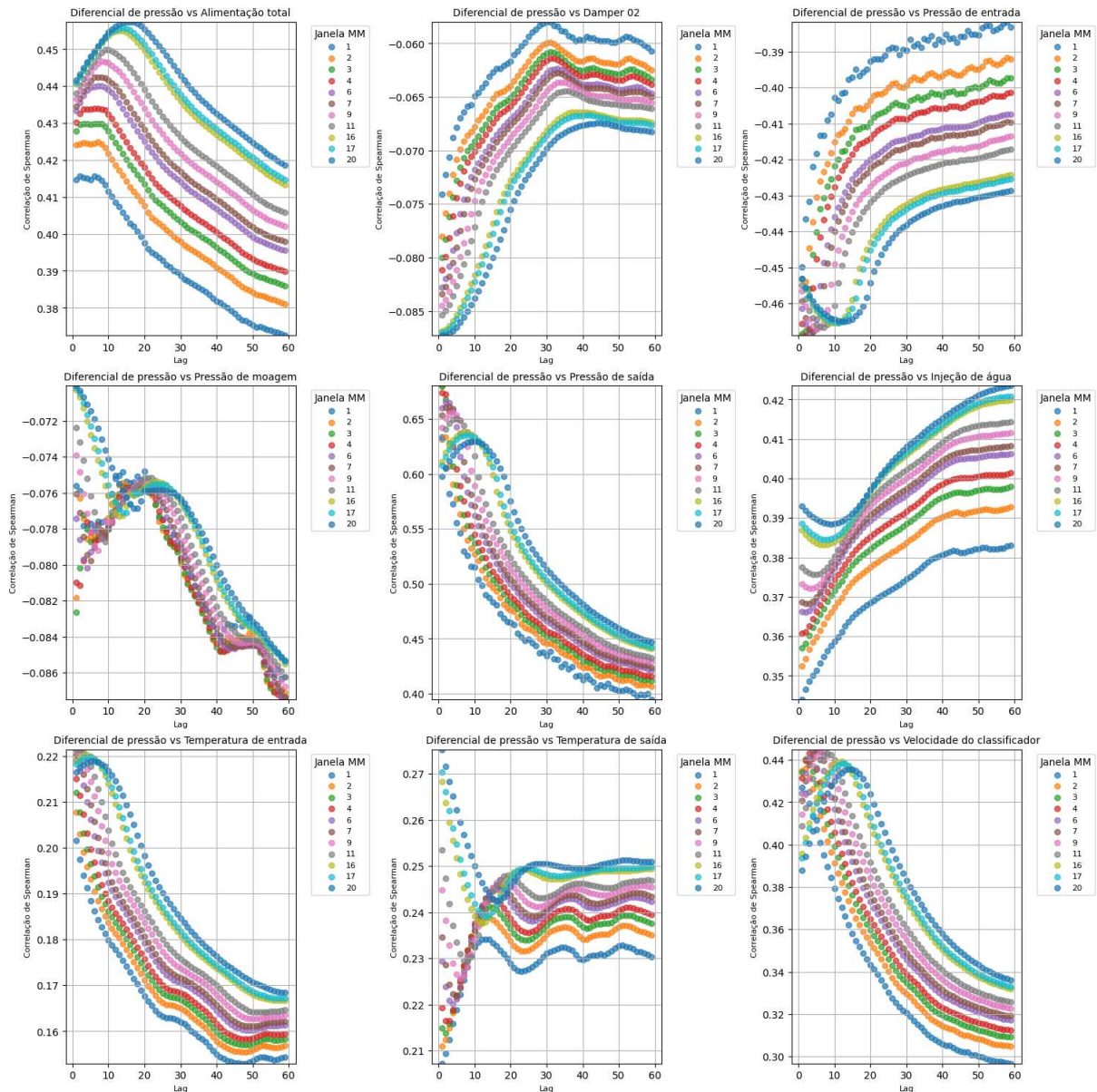


Fonte: Elaborado pela autora (2024).

de moagem, observam-se correlações mais elevadas com a velocidade do classificador e a temperatura de entrada, enquanto o exaustor apresenta correlações mais significativas com a alimentação total e também com a temperatura de entrada. Dessa forma, a relação entre essas variáveis e o diferencial de pressão torna-se menos expressiva.

A variável pressão de entrada apresenta uma forte correlação negativa, o que sugere que reduções nesta variável podem estar associadas a aumentos no diferencial de pressão. Em contrapartida, a pressão de saída exibe uma correlação positiva, indicando que aumentos na pressão de saída podem acarretar elevações no diferencial de pressão. Por outro lado, as variáveis velocidade do classificador e injeção de água revelam correlações de baixo valor, o que sugere

Figura 19 – Correlação cruzada de Spearman para diferentes valores de média móvel



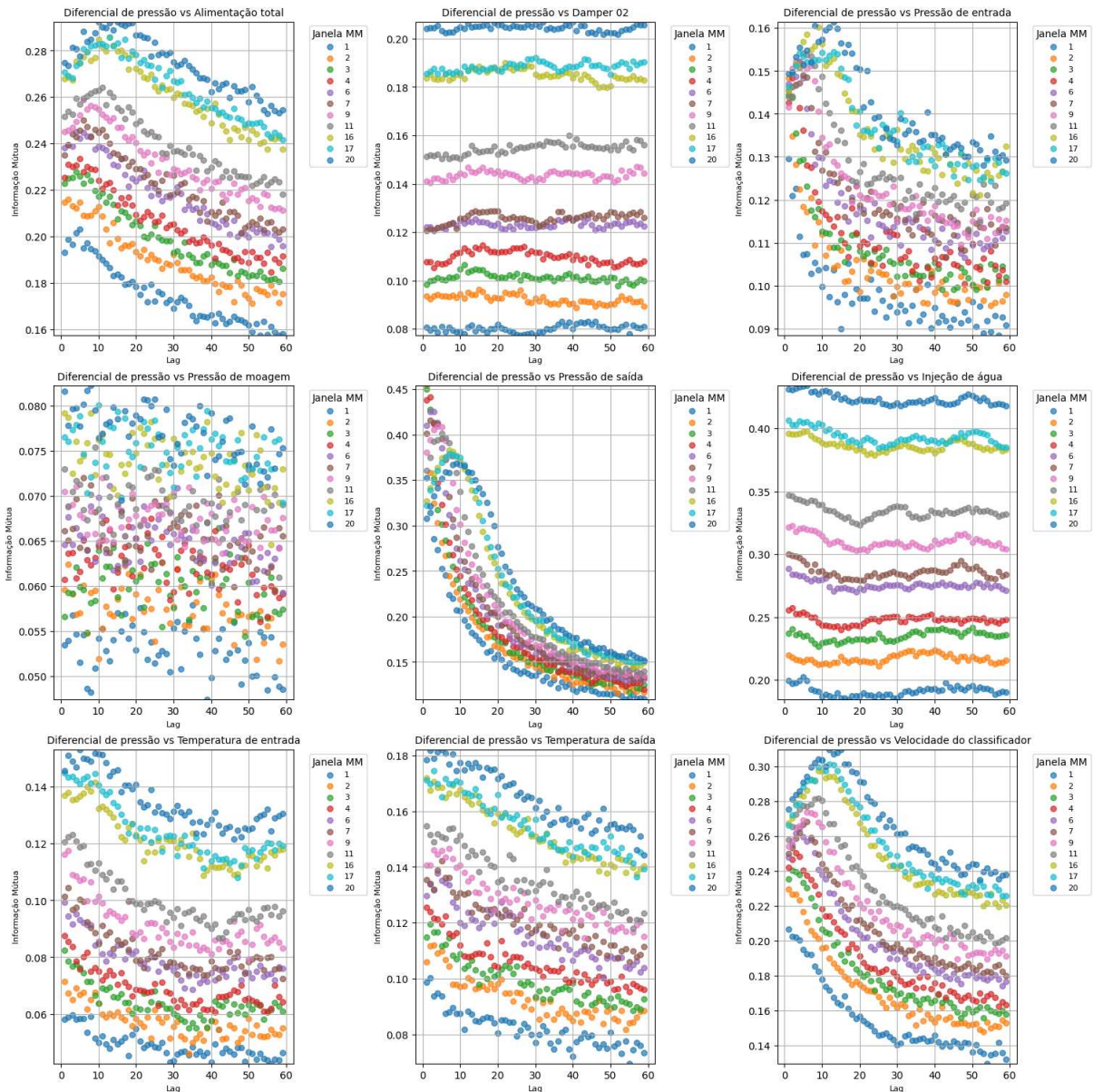
Fonte: Elaborado pela autora (2024).

que seu impacto direto sobre a variação do diferencial de pressão é relativamente limitado.

#### 6.4 Análise das relações via correlação de Spearman

A mesma abordagem, utilizando mapas de calor, foi empregada na análise da correlação de Spearman, conforme ilustrado na Figura 22, revelando padrões semelhantes aos observados na correlação de Pearson. A análise indica que a pressão de entrada apresenta uma correlação negativa com o diferencial de pressão, evidenciando uma relação inversa entre ambas. Em contrapartida, a pressão de saída exibe uma correlação positiva forte, sugerindo uma associação direta e significativa com o diferencial de pressão

Figura 20 – Correlação cruzada de Informação Mútua para diferentes valores de média móvel



Fonte: Elaborado pela autora (2024).

De maneira similar, variáveis como injeção de água, pressão de moagem e temperatura apresentam correlações fracas com o diferencial de pressão, sugerindo um impacto limitado. A pressão de moagem e o *damper O2* mostram correlações quase desprezíveis, possivelmente devido à influência predominante da temperatura de entrada, que dificulta a identificação de uma relação clara entre essas variáveis e o diferencial de pressão.

## 6.5 Análise das relações via Informação Mútua

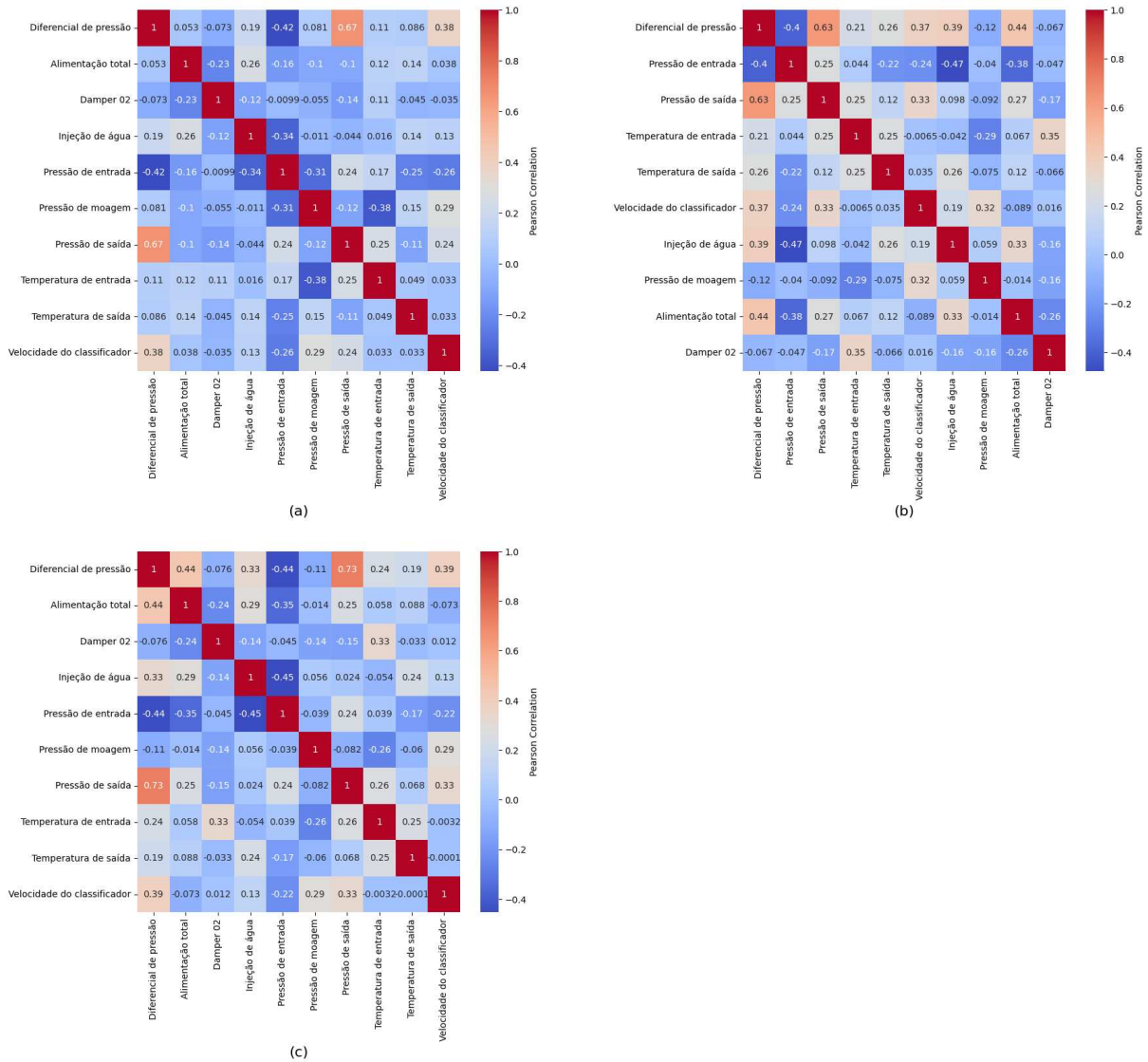
A análise da Informação Mútua demonstrou uma excelente aderência às variáveis explicativas que apresentavam correlações desprezíveis, conforme descrito nos tópicos de análise

Tabela 7 – Tabela de valores de lag e janela de média móvel selecionados para suavização de dados

Variável	Valor da Corr.	Tipo de Corr.	Lag	Janela
Alimentação total	0,4986	Pearson	6	20
Damper 02	0,2130	Informação Mútua	24	20
Injeção de água	0,4202	Pearson	30	20
Pressão de entrada	-0,4569	Pearson	1	20
Pressão de moagem	0,1595	Informação Mútua	7	20
Pressão de saída	0,7172	Pearson	1	4
Temperatura de entrada	0,2529	Pearson	1	20
Temperatura de saída	0,2896	Pearson	30	20
Velocidade do classificador	0,4504	Spearman	3	9

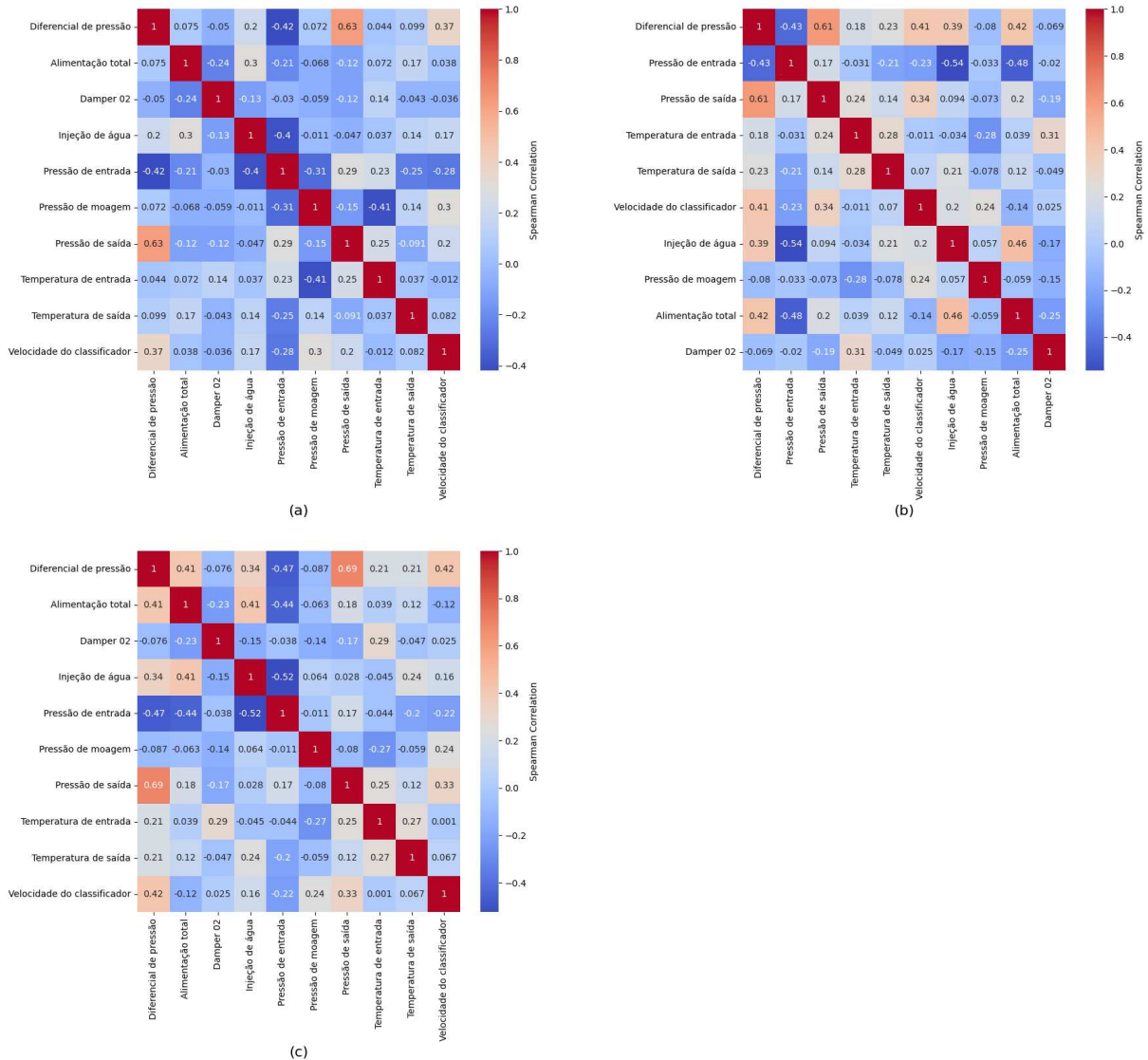
Fonte: Elaborado pela autora (2024)

Figura 21 – Matrizes *heatmap* da correlação de Pearson: (a) com os dados não tratados, (b) com os dados filtrados e (c) com a aplicação das suavizações.



Fonte: Elaborado pela autora (2024).

Figura 22 – Matrizes *heatmap* da correlação de Spearman: (a) com os dados não tratados, (b) com os dados filtrados e (c) com a aplicação das suavizações.



Fonte: Elaborado pela autora (2024).

anteriores. Esse resultado indica que tais variáveis possuem um comportamento mais complexo do que uma relação meramente linear com a variável resposta. Conforme ilustrado na Figura 23, que exibe as relações da Informação Mútua, observa-se um aprimoramento na identificação dessas interações.

A partir da análise dos mapas de calor, foi possível inferir um aumento nos valores das relações conforme os dados foram tratados. Verificou-se que a pressão de entrada, a pressão de saída e a alimentação total apresentam uma forte correlação com o diferencial de pressão, influenciando diretamente seu comportamento. Além disso, a alimentação total demonstrou uma correlação significativa com a pressão de moagem e a pressão de saída, sugerindo que um aumento na alimentação do moinho resulta em elevações nessas variáveis.

Por outro lado, a velocidade do classificador apresentou correlações mais fracas com a maioria das variáveis, sugerindo que seu impacto pode ser relativamente menor em comparação com outros parâmetros. Esse comportamento pode ser justificado por sua forte correlação com a pressão de saída do moinho, uma vez que uma redução nessa pressão indica uma maior seletividade na classificação do produto final.

Além disso, identificou-se a existência de correlações positivas perfeitas entre algumas variáveis explicativas, o que sugere que essas variáveis carregam essencialmente a mesma informação. Dessa forma, a remoção de uma delas não acarretaria perda significativa em um modelo preditivo. Dentre as variáveis analisadas, a injeção de água foi a que apresentou maior recorrência desse fenômeno.

Por fim, em termos operacionais, os valores de Informação Mútua demonstraram uma maior aderência ao processo de moagem de cimento em comparação com as demais correlações. Dessa forma, pode-se inferir que as relações entre as variáveis explicativas, bem como entre estas e o diferencial de pressão, podem apresentar um comportamento predominantemente complexo.

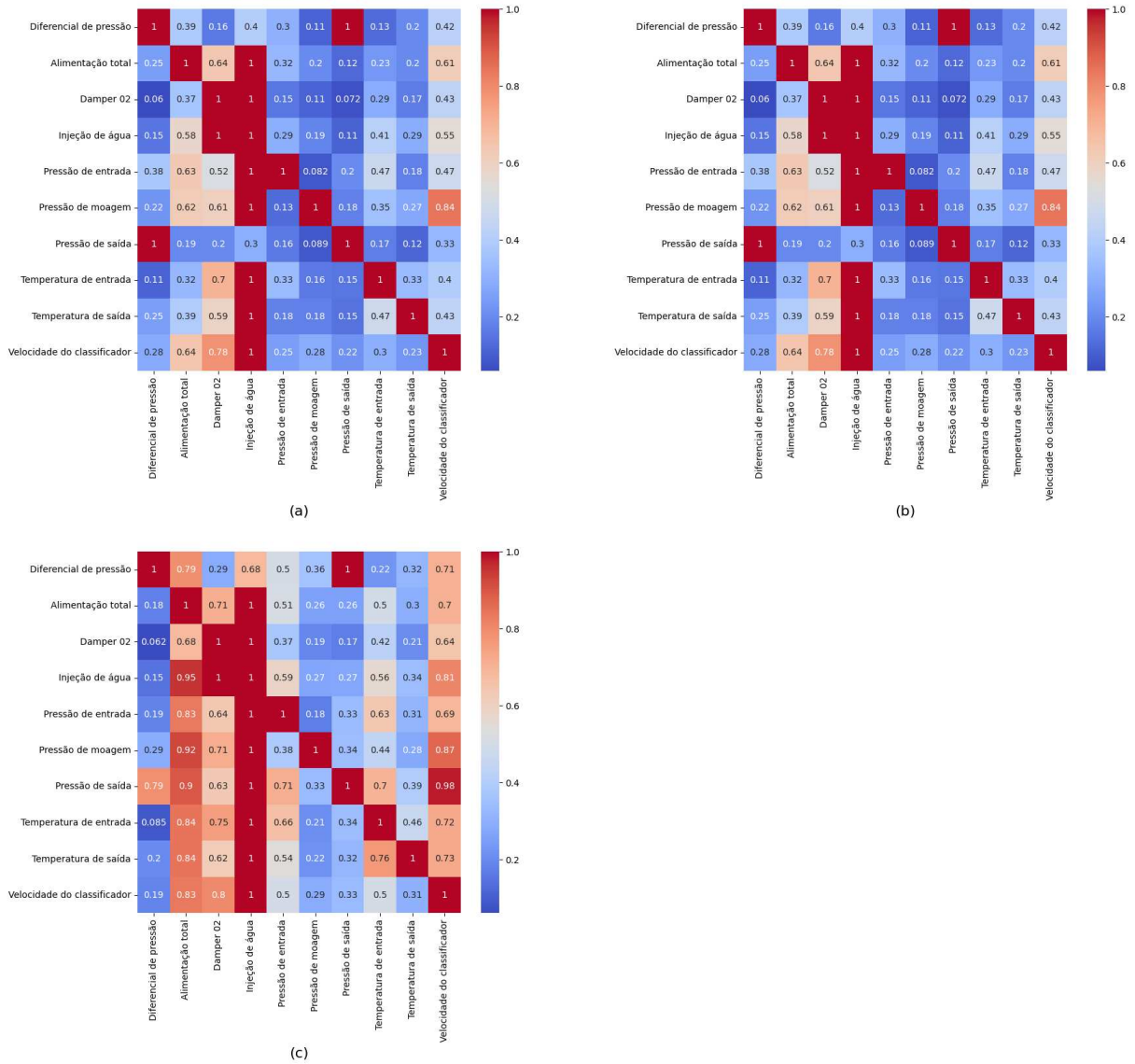
## 6.6 Predição com *Random Forest Regressor*

Inicialmente, a predição foi realizada não contemplando os parâmetros explícitos e utilizando dados com suavizações por média móvel e lags. No total, foram consideradas 105.061 observações para cada variável, das quais 80% foram alocadas para o treinamento e 20% para a etapa de teste. A Figura 24 apresenta os resultados de teste obtidos por meio dessa abordagem. Posteriormente, foi elaborado um gráfico SHAP (*SHapley Additive exPlanations*) com o objetivo de investigar o impacto individual das variáveis no modelo preditivo.

A Figura 25 ilustra o impacto das variáveis na predição do modelo. O eixo horizontal apresenta os valores SHAP, os quais indicam a influência de cada variável na saída do modelo. Um valor SHAP positivo indica que a variável contribuiu para o aumento da previsão, enquanto um valor negativo sugere que a variável reduziu a previsão. As cores em tons de azul representam valores baixos, enquanto as cores em tons de rosa correspondem a valores elevados.

Dessa forma, pode-se inferir que a pressão de saída é a variável que exerce a maior influência no modelo, apresentando variações tanto positivas quanto negativas em relação à previsão. Em seguida, destaca-se a pressão de entrada, o que não é surpreendente, dado que ambas as pressões determinam o diferencial de pressão. Em terceiro lugar, encontra-se a alimentação total, evidenciando a relação intrínseca entre essas variáveis.

Figura 23 – Matrizes *heatmap* da correlação de Informação Mútua: (a) com os dados não tratados, (b) com os dados filtrados e (c) com a aplicação das suavizações.

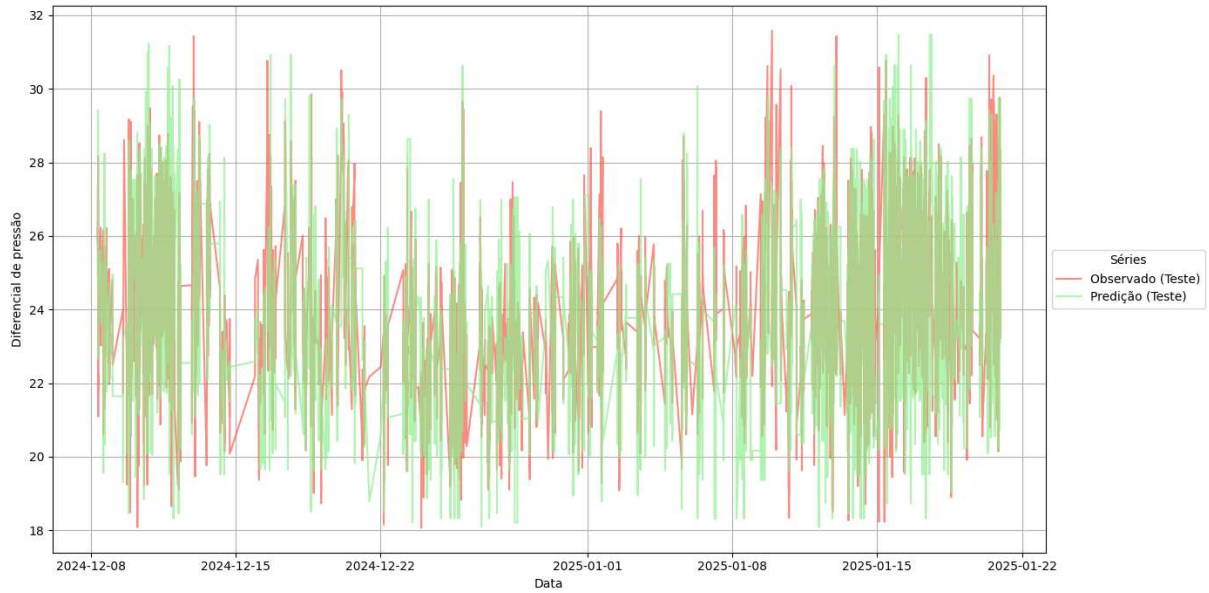


Fonte: Elaborado pela autora (2024).

O aumento da pressão no moinho está diretamente associado ao aumento da alimentação. Da mesma forma, a saída de material reflete uma maior produtividade. A injeção de água, por sua vez, revela uma influência mínima na predição do modelo, apresentando os menores intervalos de valores SHAP. No que se refere ao desempenho do modelo, a Tabela 8 apresenta os valores das métricas da predição não parametrizada.

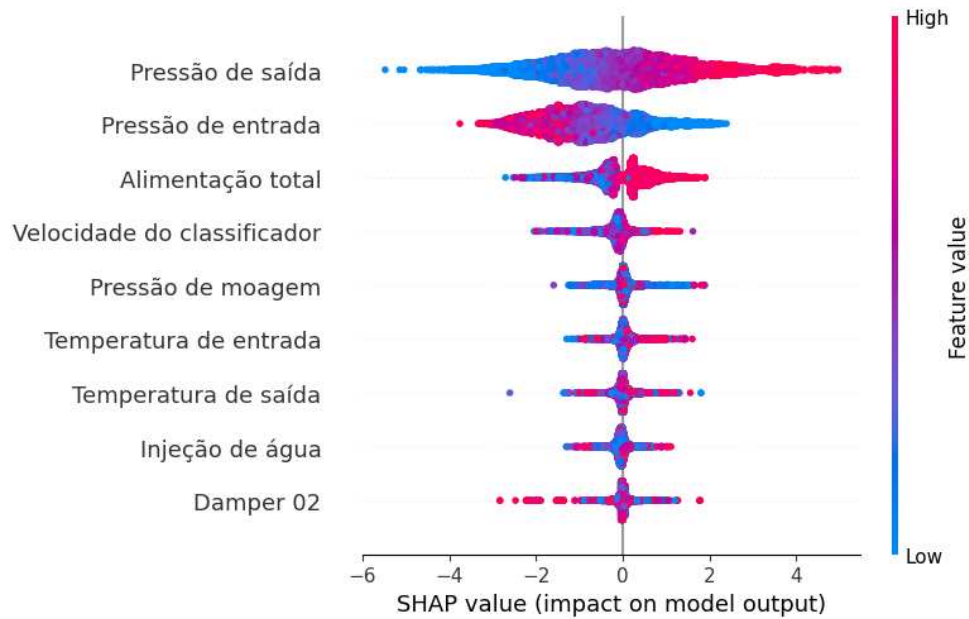
Os valores de MAE e RMSE indicam erros médios relativamente baixos, sugerindo uma boa capacidade preditiva. No entanto, o MASE, por ser superior a 1, evidencia que o modelo apresenta um desempenho inferior ao de um método de referência. Esse resultado destaca a necessidade de aprimoramentos, seja por meio do ajuste de parâmetros, do refinamento das variáveis predictoras ou da consideração de técnicas mais avançadas, a fim de elevar a precisão e

Figura 24 – Gráfico dos valores observados e preditos do diferencial de pressão nas amostras de teste.



Fonte: Elaborado pela autora (2025).

Figura 25 – Análise SHAP das variáveis que influenciam a predição do diferencial de pressão do moinho.



Fonte: Elaborado pela autora (2025).

Tabela 8 – Valores das métricas para teste da predição não parametrizada.

Métrica	Descrição	Valor
MAE	Erro absoluto médio	1,4594
RMSE	Raiz do erro quadrático médio	1,9014
MASE	Erro médio escalonado	2,3640

Fonte: Elaborado pela autora (2025)

a capacidade de generalização do modelo.

Dessa forma, os próximos subtópicos deste capítulo tratam da definição dos parâmetros de treinamento do modelo de predição com base na validação cruzada, além de apresentarem a análise do comportamento da predição com baixa variabilidade de alimentação total.

### 6.6.1 Configuração dos parâmetros do modelo

Os testes preliminares, nos quais foram combinados diferentes valores para os parâmetros analisados, indicaram que o número de árvores ( $n\_estimators$ ) e a semente para reprodutibilidade ( $random\_state$ ) são os fatores de maior impacto nas métricas de desempenho do modelo preditivo em estudo. Com base nesse resultado, foram definidos valores específicos para esses parâmetros, visando avaliar suas combinações e identificar a configuração que proporciona o melhor desempenho do modelo.

A seleção dos valores adotados considerou a relevância de cada parâmetro no contexto da análise incremental. Observou-se, por exemplo, que a métrica de desempenho apresentou variações mais expressivas quando o número de árvores foi aumentado em 30 unidades. Assim, os valores estabelecidos para esse parâmetro foram 10, 40, 70 e 100. De maneira análoga, para a semente de reprodutibilidade, adotaram-se os valores 1, 42, 84 e 126.

A análise foi conduzida considerando todas as combinações possíveis desses valores, totalizando 16 execuções. A Tabela 9 apresenta as combinações  $C$  adotadas no estudo. Para a validação cruzada, utilizou-se uma divisão em 10 *folds*. Em cada execução, as métricas de desempenho foram calculadas para cada grupo selecionado, conforme ilustrado nas Figuras 26, 27 e 28.

A partir dessa análise, observou-se que a combinação  $C14$  apresentou os menores valores nas métricas avaliadas. Com base nesses resultados, optou-se por treinar novamente o modelo utilizando a base de dados, empregando 100 árvores de decisão ( $n\_estimators = 100$ ) e 42 sementes ( $random\_state = 42$ ), com o objetivo de obter as métricas de desempenho. A Figura 29 apresenta o gráfico com os valores observados e de teste para a predição com parametrização ajustada.

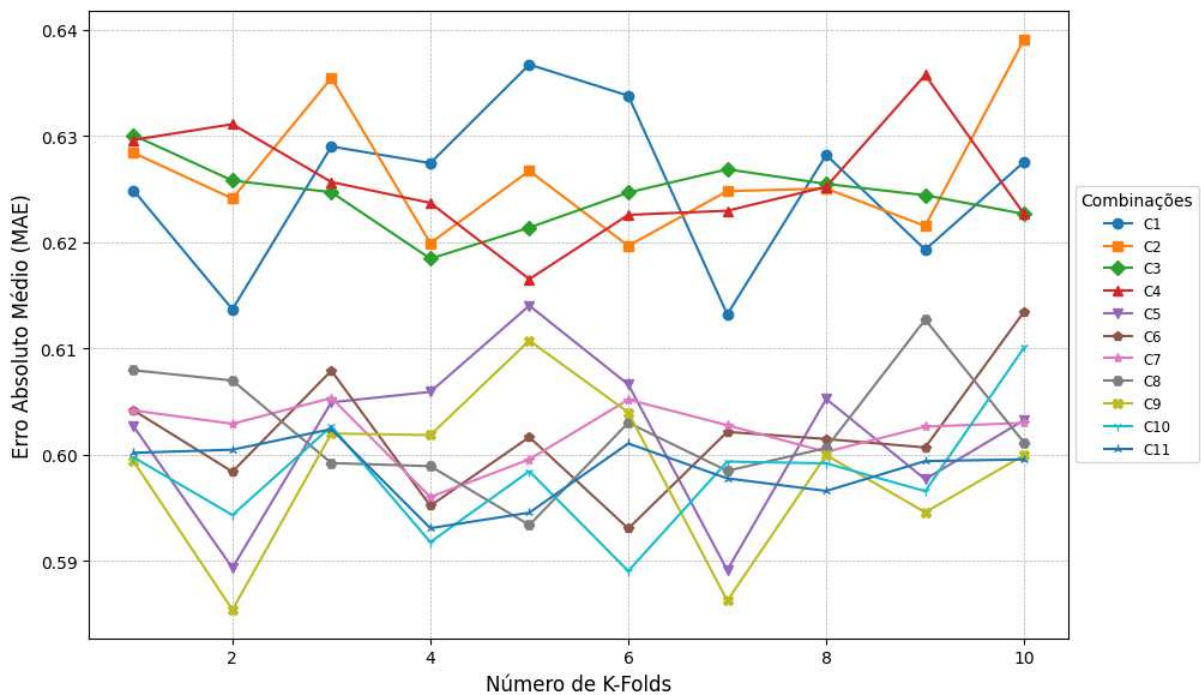
A Tabela 10 apresenta os valores das métricas para o teste da predição, considerando a aplicação dos parâmetros anteriormente mencionados. Observa-se uma redução nos valores das métricas em comparação aos resultados obtidos na etapa anterior, o que evidencia a melhoria na capacidade do modelo em gerar previsões mais precisas e consistentes.

Tabela 9 – Combinação de valores de parâmetros para validação cruzada

Combinação	n_estimator	random_state
C1	10	1
C2	10	42
C3	10	84
C4	10	126
C5	40	1
C6	40	42
C7	40	84
C8	40	126
C9	70	1
C10	70	42
C11	70	84
C12	70	126
C13	100	1
C14	100	42
C15	100	84
C16	100	126

Fonte: Elaborado pela autora (2025)

Figura 26 – Gráfico da evolução do erro absoluto médio dos grupos para cada combinação de parâmetros

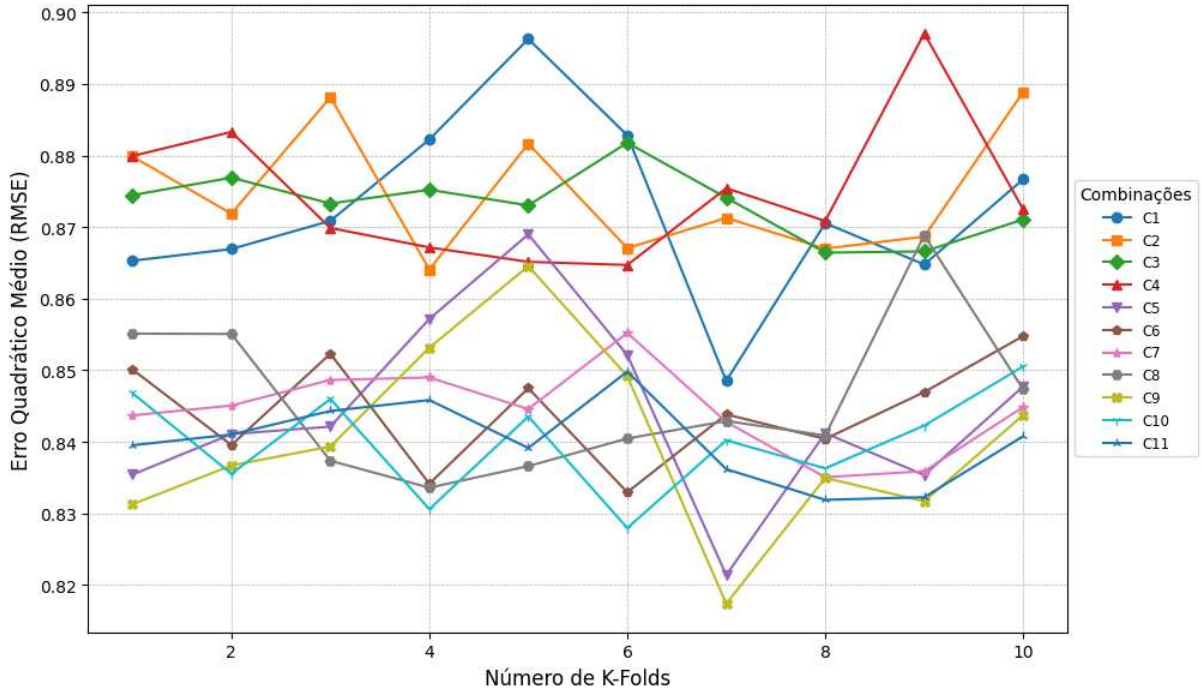


Fonte: Elaborado pela autora (2025).

### 6.6.2 Predição com baixa variabilidade de pressão de saída

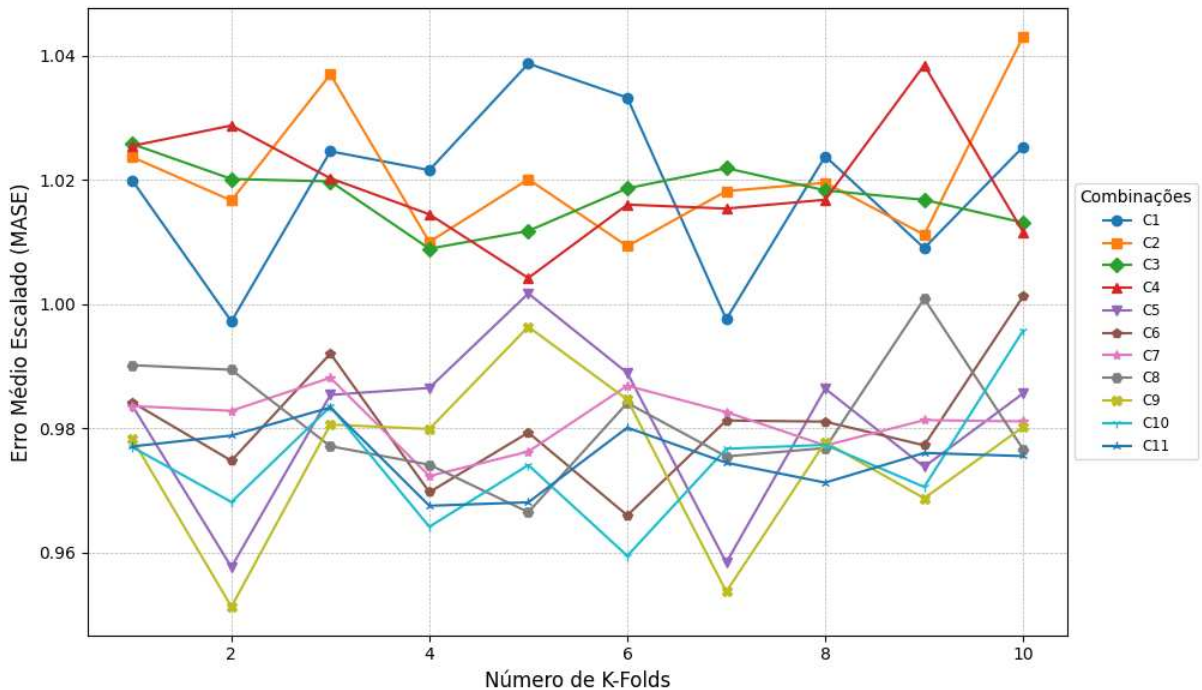
A pressão de saída é um parâmetro operacional relevante, diretamente associado ao diferencial de pressão do moinho, pois reflete a resistência encontrada pelo fluxo de ar ao

Figura 27 – Gráfico da evolução do erro quadrático médio da raiz dos dos grupos para cada combinação de parâmetros



Fonte: Elaborado pela autora (2025).

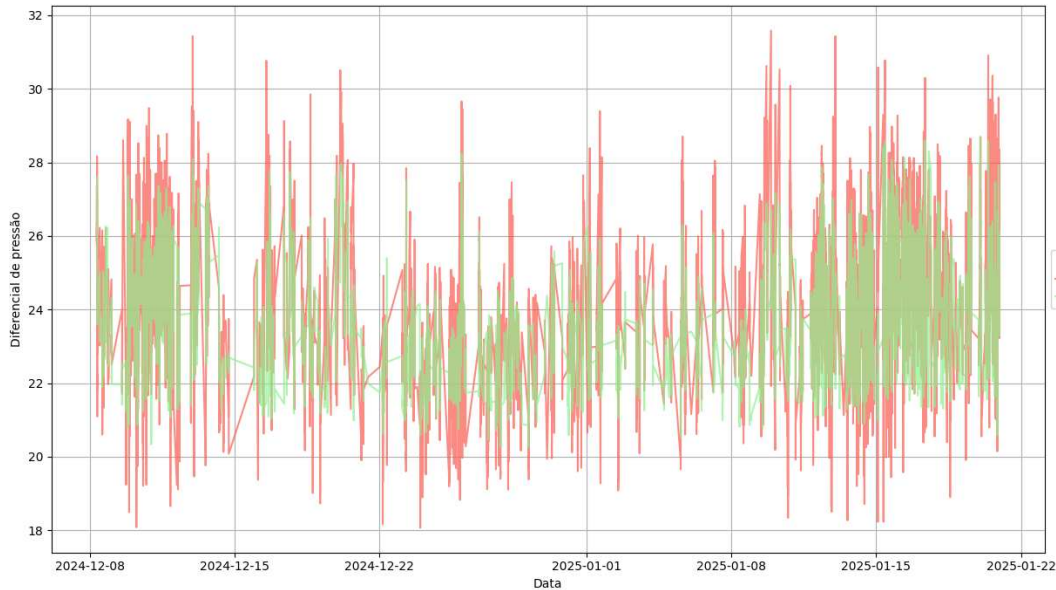
Figura 28 – Gráfico da evolução do erro absoluto médio escalonado dos grupos para cada combinação de parâmetros



Fonte: Elaborado pela autora (2025).

sair do sistema. Quanto maior a pressão de saída, maior o diferencial de pressão, o que indica um aumento na alimentação de matérias-primas. Diante disso, este estudo teve como objetivo avaliar o desempenho da predição para um intervalo específico de pressão de saída. O intervalo

Figura 29 – Gráfico dos valores observados e preditos do diferencial de pressão nas amostras de teste, com a parametrização ajustada.



Fonte: Elaborado pela autora (2025).

Tabela 10 – Valores das métricas para teste da predição parametrizada.

Métrica	Descrição	Valor
MAE	Erro absoluto médio	0,9872
RMSE	Raiz do erro quadrático médio	1,2958
MASE	Erro médio escalonado	1,5991

Fonte: Elaborado pela autora (2025)

selecionado, de 30 a 36 t/h, foi escolhido devido à maior disponibilidade de dados para análise.

As Tabelas 11, 12 e 13 apresentam os valores de janela média e lag para as correlações máximas de Pearson, Spearman e Informação Mútua, respectivamente. Para os dados com baixa variabilidade de pressão de saída, as correlações de Pearson e Spearman indicam, em sua maioria, relações fracas ou muito fracas. No entanto, a pressão de entrada se destaca com uma correlação moderada e negativa, sugerindo que seu aumento tende a reduzir a pressão de saída.

Além disso, a velocidade do classificador apresenta uma correlação classificada como forte na análise de Spearman, indicando uma possível relação não linear mais significativa. Já na análise de Informação Mútua, observa-se uma dependência fraca, mas relevante, entre a pressão de saída e a alimentação total, reforçando a influência dessa variável na dinâmica do sistema.

A correlação de Informação Mútua apresenta resultados mais expressivos para o *damper 02*, pressão de entrada e pressão de moagem. Em relação ao lag e à janela, observa-se que os valores para as correlações de Pearson e Spearman tendem a ser mais elevados nas

variáveis com maior impacto, refletindo a complexidade das interações entre esses parâmetros e o diferencial de pressão. Isso sugere que, mesmo sob baixa variabilidade, a influência dessas variáveis pode ocorrer de forma defasada no tempo.

Tabela 11 – Tabela dos valores de lag e da janela de média móvel para os valores máximos de correlação de Pearson, referentes a dados com baixa variabilidade da pressão de saída.

Variável	Corr. Pearson	Classificação da correlação	Lag	Janela
Alimentação total	0,4695	Fraca	6	20
Damper 02	-0,0561	Desprezível	1	20
Injeção de água	0,4511	Fraca	30	20
Pressão de entrada	-0,5827	Moderada	1	20
Pressão de moagem	-0,1045	Muito fraca	30	20
Pressão de saída	0,6266	Moderada	1	6
Temperatura de entrada	0,2445	Muito fraca	1	20
Temperatura de saída	0,3212	Muito fraca	1	20
Velocidade do classificador	0,4027	Fraca	4	15

Fonte: Elaborado pela autora (2025)

Tabela 12 – Tabela dos valores de lag e da janela de média móvel para os valores máximos de correlação de Spearman, referentes a dados com baixa variabilidade de pressão de saída.

Variável	Corr. Spearman	Classificação da correlação	Lag	Janela
Alimentação total	0,4387	Muito fraca	7	20
Damper 02	-0,0630	Desprezível	1	20
Injeção de água	0,4524	Fraca	30	20
Pressão de entrada	-0,5878	Fraca	1	20
Pressão de moagem	-0,0711	Muito fraca	30	1
Pressão de saída	0,6139	Moderada	1	5
Temperatura de entrada	0,2080	Desprezível	1	14
Temperatura de saída	0,2875	Desprezível	1	20
Velocidade do classificador	0,4238	Forte	5	18

Fonte: Elaborado pela autora (2025)

A correlação de Pearson, por sua vez, demonstrou maior sensibilidade às variações das variáveis analisadas, especialmente em relação à alimentação total. Esse comportamento sugere que essa métrica responde de maneira mais acentuada a pequenas oscilações na pressão de saída, resultando em valores mais elevados em comparação às demais correlações. Os resultados desta análise estão apresentados na Tabela 14.

No que tange às métricas de predição, a modelagem sob condições de baixa variabilidade da pressão de saída apresentou valores inferiores para os indicadores de desempenho. Esse resultado sugere que a avaliação do modelo sob tais condições é mais adequada. Ademais,

Tabela 13 – Tabela dos valores de lag e da janela de média móvel para os valores máximos de Informação Mútua, referentes a dados com baixa variabilidade da pressão de saída.

Variável	Informação Mútua	Classificação da correlação	Lag	Janela
Alimentação total	0,4229	Fraca	3	20
Damper 02	0,2050	Muito fraca	13	20
Injeção de água	0,4044	Fraca	3	20
Pressão de entrada	0,4000	Fraca	1	20
Pressão de moagem	0,1376	Muito fraca	7	20
Pressão de saída	0,4265	Fraca	1	20
Temperatura de entrada	0,2222	Muito fraca	2	20
Temperatura de saída	0,1967	Muito fraca	2	20
Velocidade do classificador	0,3839	Fraca	4	20

Fonte: Elaborado pela autora (2025)

Tabela 14 – Tabela de valores de lag e janela de média móvel para dados com baixa variabilidade de alimentação total

Variável	Valor da Corr.	Tipo de Corr.	Lag	Janela
Alimentação total	0,4695	Pearson	6	20
Damper 02	0,2050	Informação Mútua	13	20
Injeção de água	0,4524	Spearman	30	20
Pressão de entrada	0,4000	Informação Mútua	1	20
Pressão de moagem	0,1376	Informação Mútua	7	20
Pressão de saída	0,6266	Pearson	1	6
Temperatura de entrada	0,2445	Pearson	1	20
Temperatura de saída	0,3212	Pearson	1	20
Velocidade do classificador	0,4238	Spearman	5	18

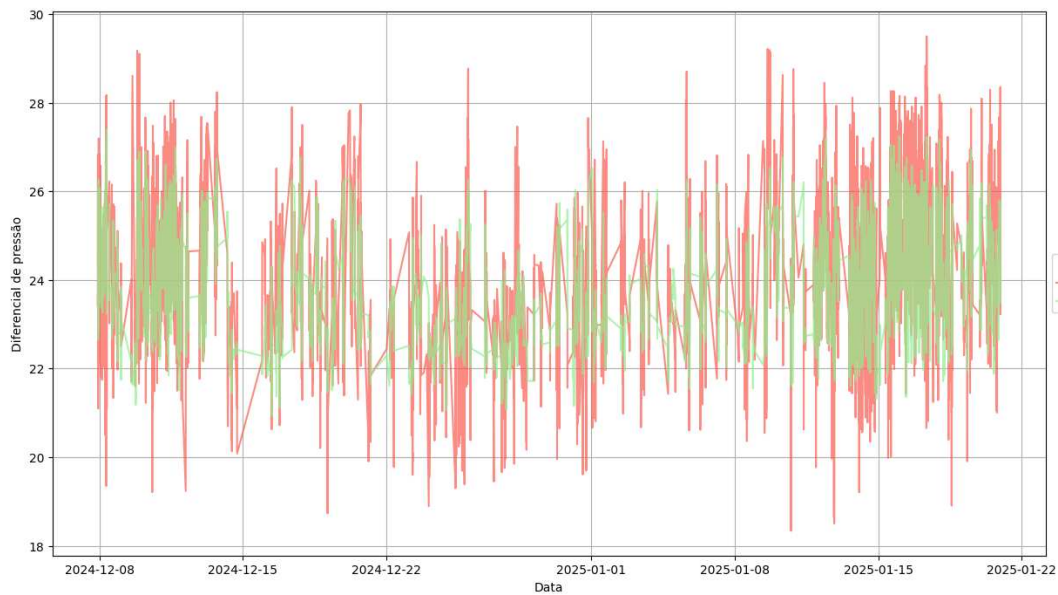
Fonte: Elaborado pela autora (2025)

a Figura 30 ilustra o comportamento temporal da predição dessa variável, possibilitando uma análise visual da aderência do modelo aos dados de teste.

Tabela 15 – Métricas de desempenho do modelo de predição com baixa variabilidade de pressão de saída

Métrica	Descrição	Valor
MAE	Erro absoluto médio	0,9438
RMSE	Raiz do erro quadrático médio	1,2089
MASE	Erro médio escalonado	1,6023

Figura 30 – Gráfico dos valores observados e preditos do diferencial de pressão na amostra de teste para os dados com baixa variabilidade de pressão de saída.



Fonte: Elaborado pela autora (2025).

## 7 CONSIDERAÇÕES FINAIS

O presente estudo investigou as variáveis operacionais que influenciam a pressão diferencial em um moinho de rolos verticais utilizado na produção de cimento Portland. A relevância dessa pesquisa está na importância de manter a pressão diferencial sob controle operacional ideal, visando evitar a instabilidade do equipamento, o que pode acarretar aumento nos custos de manutenção e redução na produtividade.

Foram realizadas análises estatísticas, juntamente com a modelagem preditiva utilizando o modelo *Random Forest Regressor*, para prever o comportamento do diferencial de pressão com base em dados históricos de uma unidade cimenteira. Os resultados obtidos indicaram correlações significativas entre o diferencial de pressão e outras variáveis operacionais, fornecendo insights valiosos para a otimização dos processos de moagem.

Destacou-se, particularmente, que a pressão de entrada e saída, bem como a quantidade total de alimentação, apresentam fortes correlações com a pressão diferencial. Uma correlação negativa foi observada entre a pressão de entrada e a pressão diferencial, sugerindo que aumentos na pressão de entrada resultam em redução da pressão diferencial. Por outro lado, a pressão de saída apresentou uma correlação positiva, indicando que pressões de saída mais elevadas resultam em maiores níveis de pressão diferencial.

A análise de defasagem temporal foi crucial para compreender a relação entre as variáveis explicativas e o diferencial de pressão ao longo do tempo. A definição do lag e da janela de média móvel mostrou-se eficaz na identificação das melhores configurações de processamento de dados, sendo que a correlação de Pearson apresentou os maiores valores para a maioria das variáveis analisadas. As variáveis como a pressão de saída, a alimentação total e a velocidade do classificador demonstraram correlações significativas, tanto lineares quanto não lineares, com o diferencial de pressão.

As correlações de Spearman confirmaram os achados das correlações de Pearson. A aplicação da Informação Mútua, no entanto, destacou a complexidade das interações, revelando correlações mais fortes entre variáveis previamente identificadas com relações fracas. Além disso, indicou correlações positivas perfeitas entre algumas variáveis, sugerindo redundância informacional e a oportunidade de simplificar os modelos preditivos sem comprometer seu desempenho. Esses resultados ressaltam a necessidade de cautela no processamento dessas variáveis durante a modelagem, destacando a importância de investigações mais detalhadas.

A aplicação de técnicas de aprendizado de máquina demonstrou um considerável

potencial preditivo, com o modelo apresentando desempenho aprimorado após o ajuste de parâmetros. No entanto, o erro absoluto médio indicou que o modelo ainda requer refinamento adicional para superar um desempenho básico. Os resultados sugerem que, em condições de baixa variabilidade da pressão de saída, o modelo alcança maior precisão, sendo sua performance passível de otimização adicional por meio do ajuste das variáveis mais influentes.

## REFERÊNCIAS

- ABCP. **História do cimento**. 2024. Acesso em: 30 nov. 2024. Disponível em: <https://abcp.org.br/cimento/>.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 16687**: Cimento portland – requisitos. Rio de Janeiro, 2018.
- BELMAJDOUB, S. A. F. Efficient machine learning model to predict fineness, in a vertical raw meal of morocco cement plant. **Cement and Concrete Research**, Elsevier, v. 121, p. 211–217, 2019. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0921883119304583>.
- BISHOP, C. M. **Pattern recognition and machine learning**. New York : Springer Science + Business Media, LLC, 2006.
- BREIMAN, L. Bagging predictors. **Machine Learning**, Springer, v. 24, n. 2, p. 123–140, 1996.
- BUSSAB, W. de O.; MORETTIN, P. A. **Estatística Básica**. 6. ed. São Paulo: Saraiva, 2010. ISBN 978-85-02-08177-2.
- COSTA, E. B. da; NOBRE, T. R. S.; GUERREIRO, A. Q.; MANCIO, M.; KIRCHHEIM, A. P. Clínquer portland com reduzido impacto ambiental. **Ambiente Construído**, v. 13, n. 2, p. 75–86, abr./jun. 2013. ISSN 1678-8621. Disponível em: <https://www.scielo.br/j/ac/a/RQWhYNNvCNYDZs7sgvZvydB/>.
- COVER, T. M.; THOMAS, J. A. **Elements of Information Theory**. 2. ed. New York: John Wiley Sons, Inc., 2006. ISBN 0-471-06259-6.
- GÉRON, A. **Mãos à Obra: Aprendizado de Máquina com Scikit-Learn TensorFlow: Conceitos, ferramentas e técnicas para a construção de sistemas inteligentes**. 1. ed. Rio de Janeiro: Starlin Alta Editora e Consultoria Eireli, 2019. ISBN 978-85-508-0902-1.
- HEWLETT, M. L. P. **Lea’s Chemistry of Cement and Concrete**. 5. ed. Oxford: Butterworth-Heinemann, 2019.
- IBGE. **Cimento Portland: Produção**. 2023. Acesso em: 30 nov. 2024. Disponível em: <https://anuario.ibge.gov.br/2023/industria/extrativa-mineral-e-de-transformacao/aeb-2023-tabelas-extrativa-mineral-e-de-transformacao/22183-cimento-portland-producao.html>.
- ISAIA, G. C. **Materiais de Construção Civil e Princípios de Ciência e Engenharia de Materiais**. 3. ed. São Paulo: IBRACON, 2017. v. 1.
- JOHNSON, R. A.; WICHERN, D. W. **Applied Multivariate Statistical Analysis**. 6. ed. [S. l.]: Pearson Prentice Hall, 2007. ISBN 0-13-187715-1.
- LIMA, A. B. de.
- O Processo Produtivo do Cimento Portland** – Universidade Federal de Minas Gerais, Belo Horizonte, 2011.
- MEHTA, P. K.; MONTEIRO, P. J. M. **Concreto: Microestrutura, Propriedades e Materiais**. São Paulo: IBRACON, 2006.

MENESES, L. C. de.

**Estudo de variáveis que afetam a vibração do corpo de um moinho vertical de rolos** – Universidade Federal do Ceará, Fortaleza, 2023.

MITCHELL, T. M. **Machine Learning**. 1. ed. New York: McGraw-Hill, 1997.

MOORE, D. S.; MCCABE, G. **Introduction to the practice of statistics**. 6. ed. United States of America: W. H. Freeman and Company, 2009. ISBN 1-4292-1621-2.

NEVILLE, A. M. **Properties of Concrete**. 5th. ed. London: Pearson, 2011. ISBN 978-0273755807.

PFEIRER, G. **Moinhos Pfeirer para a Indústria de Cimento: Potentes. comprovados. confiáveis.** Gebr. Pfeirer, 1940. Disponível em: [https://www.gebr-pfeiffer.com/fileadmin/user\\_upload/Zement-Broschuere\\_PT\\_Web.pdf](https://www.gebr-pfeiffer.com/fileadmin/user_upload/Zement-Broschuere_PT_Web.pdf).

REICHERT, M.; GEROLD, C.; FREDRIKSSON, A.; ADOLFSSON, G.; LIEBERWIRTH, H. Research of iron ore grinding in a vertical-roller-mill. **Minerals Engineering**, v. 73, p. 109–115, mar. 2015.

SCIKIT-LEARN. **Cross-validation: evaluating estimator performance**. 2024. Acesso em: 30 nov. 2024. Disponível em: [https://scikit-learn.org/stable/modules/cross\\_validation.html](https://scikit-learn.org/stable/modules/cross_validation.html).

SOUSA, C. A. S.; OLIVEIRA, S. E. A. de; NASCIMENTO, F. G. da S.; FIGUEREDO, A. S.; SILVA, E. B. da; SANTOS, P. H. M. de S.; NOLASCO, L. T. d. S. P. F. A. C. P.; BRITO, D. R. do N. Análise da utilização da escória no processo produtivo do cimento. **Revistaft**, v. 28, n. 137, 2024. Disponível em: <https://revistaft.com.br/analise-da-utilizacao-da-escoria-no-processo-produtivo-do-cimento/>.

SOUZA, H. M.; JR, A. P. S.; JR, E. F. C.; COSTA, A. O. S. Predição da composição do clínquer industrial utilizando minimização da energia livre de gibbs. **Cerâmica** 61, v. 66, n. 362, p. 56–64, 2020. Disponível em: <https://www.scielo.br/j/ce/a/W57CGTHgjrmyCwDwB8KJwWN/>.

SPEARMAN, C. The proof and measurement of association between two things. **The American Journal of Psychology**, v. 15, n. 1, p. 72–101, jan 1904.

TOHRY S. CHEHREH CHELGANI, S. M. M. N. A. Efficient machine learning model to predict fineness, in a vertical raw meal of morocco cement plant. **Materials Today: Proceedings**, Elsevier, v. 53, p. 9296–9301, 2022. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2590123022005035>.