

UNIVERSIDADE FEDERAL DE ALAGOAS
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM RECURSOS HÍDRICOS E SANEAMENTO



RAIRA SILVA BARBOSA

**EFEITO DAS VARIÁVEIS COMPOSICIONAIS E OPERACIONAIS NA
PRODUÇÃO DE HIDROGÊNIO A PARTIR DA DIGESTÃO ANAERÓBIA DE
RESÍDUOS AGROINDUSTRIAIS**

Maceió

2024

RAIRA SILVA BARBOSA

**EFEITO DAS VARIÁVEIS COMPOSICIONAIS E OPERACIONAIS NA
PRODUÇÃO DE HIDROGÊNIO A PARTIR DA DIGESTÃO ANAERÓBIA DE
RESÍDUOS AGROINDUSTRIAIS**

Dissertação apresentado ao Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Centro de Tecnologia da Universidade Federal de Alagoas, como requisito parcial para obtenção do título de Mestre em Recursos Hídricos e Saneamento

Orientador: Prof. Dr. Eduardo Lucena
Cavalcante de Amorim

Maceió

2024

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico

Bibliotecária: Helena Cristina Pimentel do Vale CRB-4/661

B238e Barbosa, Raira Silva.

Efeito das variáveis composicionais e operacionais na produção de hidrogênio a partir da digestão anaeróbia de resíduos agroindustriais / Raira Silva Barbosa. – 2025.

129 f. : il.

Orientador: Eduardo Lucena Cavalcante de Amorim.

Dissertação (mestrado em Recursos Hídricos e Saneamento) – Universidade Federal de Alagoas. Centro de Tecnologia, Maceió, 2025.

Bibliografia: f. 108-129.

1. Produção de hidrogênio. 2. Resíduos agroindustriais. 3. Digestão anaeróbia.
4. Análise de componentes principais. 5. Análise de regressão. I. Título.

CDU: 662.7



Folha de Aprovação

RAIRA SILVA BARBOSA

EFEITO DAS VARIÁVEIS COMPOSICIONAIS E OPERACIONAIS NA PRODUÇÃO DE HIDROGÊNIO A PARTIR DA DIGESTÃO ANAERÓBIA DE RESÍDUOS AGROINDUSTRIAIS

Dissertação apresentada ao Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Centro de Tecnologia da Universidade Federal de Alagoas, como requisito para obtenção do título de Mestre em Recursos Hídricos e Saneamento.

Em: 24 /01 /2025

Documento assinado digitalmente
gov.br EDUARDO LUCENA CAVALCANTE DE AMORIM
Data: 02/02/2025 08:20:31-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Eduardo Lucena Cavalcante de Amorim
(Orientador - PPGRHS/ CTEC/UFAL)

Documento assinado digitalmente
gov.br FERNANDA SANTANA PEITER
Data: 03/02/2025 09:56:58-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. Fernanda Santana Peiter
(Coorientadora - UFAL)

Banca examinadora:

Documento assinado digitalmente
gov.br KARINA RIBEIRO SALOMON
Data: 04/02/2025 11:53:41-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. Karina Ribeiro Salomon
(Examinadora Interna – PPGRHS/ CTEC/UFAL)

Documento assinado digitalmente
gov.br NATALIA DE SOUZA PELINSON
Data: 31/01/2025 22:24:41-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. Natália de Souza Pelinson
(Examinadora Externa à Instituição - UFSCAR)

AGRADECIMENTOS

A jornada acadêmica é repleta de desafios, aprendizados e momentos de superação. Ao chegar ao fim deste ciclo, é impossível não reconhecer e agradecer às pessoas que tornaram essa trajetória possível.

Primeiramente, quero agradecer ao Prof. Dr. Eduardo Lucena Cavalcante de Amorim, meu orientador, pela orientação precisa, pela paciência e pelo constante incentivo ao longo deste trabalho. Seu comprometimento com a pesquisa e sua dedicação foram fundamentais para meu crescimento acadêmico e profissional. Da mesma forma, sou imensamente grato à Profa. Dra. Fernanda Santana Peiter, minha coorientadora, cujo apoio e contribuições enriquecedoras foram indispensáveis para o sucesso desta dissertação.

Agradeço também à minha família, pelo amor incondicional e pela base sólida que me sustentou em cada etapa deste caminho. Sem seu apoio, incentivo e compreensão nos momentos mais desafiadores, esta conquista não teria sido possível. Aos amigos que estiveram ao meu lado, compartilhando desafios, alegrias e palavras de encorajamento, deixo meu mais sincero agradecimento. A companhia de vocês tornou essa caminhada mais leve e significativa.

Por fim, sou grato a todos que, direta ou indiretamente, contribuíram para a realização deste trabalho. Cada troca de conhecimento, cada palavra de apoio e cada incentivo foram fundamentais para que eu chegasse até aqui.

RESUMO

Esta pesquisa visa analisar o potencial de resíduos agroindustriais na produção de hidrogênio por meio de digestão anaeróbia, utilizando estatísticas multivariadas para compreender as interações entre variáveis operacionais e composicionais. A partir de uma revisão sistemática de 12 estudos indexados nas bases Scopus e Sucupira, realizados por pesquisadores do Centro de Tecnologia da UFAL, identificou-se uma série de fatores para otimização da produção de hidrogênio. As análises, realizadas através do software R Project for Statistical Computing, envolveram Análise de Componentes Principais (PCA) e modelos de regressão múltipla (linear, não linear com transformação de Box-Cox, e com interações de segunda ordem), separadas por reatores operados em batelada e reatores de fluxo contínuo (RALF). Por meio da PCA, foi possível identificar componentes que explicam a variabilidade dos dados, destacando para os reatores em batelada, os fatores relacionados à carga orgânica e à composição inicial do substrato, e para os reatores contínuos, variáveis como pH inicial, tempo de detenção hidráulica (TDH), taxa de carga orgânica (TCO) e temperatura. A análise de regressão linear forneceu uma compreensão inicial das relações diretas, mas modelos não lineares e com interações de segunda ordem revelaram-se mais eficazes ao capturar a complexidade do sistema, especialmente em reatores de fluxo contínuo. O modelo com interações de segunda ordem apresentou o melhor ajuste, R^2 (0,68) e R^2 ajustado (0,58), indicando uma maior capacidade de previsão. Os resultados destacam que a importância das interações entre variáveis para otimizar o rendimento da geração de hidrogênio por processos de digestão anaeróbia a partir de resíduos agroindustriais.

Palavras-chave: Produção de hidrogênio; Resíduos agroindustriais; Digestão anaeróbia; Análise de Componentes Principais; Regressão.

ABSTRACT

This research aims to analyze the potential of agro-industrial residues in hydrogen production through anaerobic digestion, utilizing multivariate statistics to understand the interactions between operational and compositional variables that influence the process. Based on a systematic review of 12 studies indexed in the Scopus and Sucupira databases, conducted by researchers from the UFAL Technology Center, various factors were identified for optimizing hydrogen production. Analyses, performed with R Project software, included Principal Component Analysis (PCA) and multiple regression models (linear, non-linear with Box-Cox transformation, and second-order interaction models), categorized by batch and continuous reactors (RALF). The PCA identified principal components explaining data variability, highlighting, for batch reactors, components related to organic load and initial substrate composition. For continuous reactors, variables such as initial pH, hydraulic retention time (HRT), organic loading rate (OLR), and temperature were shown to play significant roles in reactor efficiency. Linear regression provided an initial understanding of direct relationships, but non-linear and second-order interaction models proved more effective in capturing system complexity, especially in continuous reactors. The second-order interaction model demonstrated the best fit, evidenced by higher R^2 (0.68) and adjusted R^2 (0.58), along with enhanced predictive capacity. Results indicate that variable interaction is crucial for maximizing hydrogen yield, providing a foundation for optimizing anaerobic digestion processes using agro-industrial residues.

Keywords: Hydrogen production; Agro-industrial residues; Anaerobic digestion; Principal Component Analysis; Regression.

LISTA DE FIGURAS

Figura 1: Representação esquemática das etapas e microrganismos envolvidos no processo de digestão anaeróbia, bem como as reações de consumo de hidrogênio.	22
Figura 2: Proposta de desenvolvimento de modelo matemático.	28
Figura 3: Principais Contribuições da Regressão Linear e Não linear no contexto da digestão anaeróbia.	34
Figura 4: Etapas Da Implementação Da PCA No Software R.	38
Figura 5: Etapas da modelagem supervisionada no R.	40
Figura 6: Mapa de calor das correlações de Spearman entre as variáveis. ...	45
Figura 7: Gráfico dos dois primeiros Componentes Principais (Modo Batelada).	51
Figura 8: Gráfico de correlações (Modo Batelada).	53
Figura 9: Gráfico de Dispersão: Valores Reais vs. Valores Previstos Modelo Não Linear (Modo Batelada).	63
Figura 10: Gráfico de Dispersão: Valores Reais vs. Valores Previstos Modelo com Interações (Modo Batelada).	71
Figura 11: Mapa de calor das correlações de Spearman entre as variáveis (Modo Contínuo).	76
Figura 12: Gráfico com as cargas dos 2 primeiros fatores (Modo Contínuo). ..	84
Figura 13: Gráfico com as cargas do 1° e 3° fatores (Modo Contínuo).	85
Figura 14: Gráfico com as cargas do 1° e 4° fatores (Modo Contínuo).	85
Figura 15: Gráfico de correlações (Modo Contínuo).	87
Figura 16: Gráfico de Dispersão: Valores Reais vs. Valores Previstos Modelo Não Linear (Modo Contínuo).	96
Figura 17: Gráfico de Dispersão: Valores Reais vs. Valores Previstos Modelo com Interações (Modo Contínuo).	104

LISTA DE TABELAS

Tabela 1: Vantagens, desvantagens e principais micro-organismos estudados para produção de hidrogênio biológico	20
Tabela 2: Comparação entre as faixas de pH aplicadas e o pH ótimo obtido na produção de H ₂ em reatores anaeróbios em batelada.	23
Tabela 3: Comparação entre pré-tratamentos do inóculo aplicados na produção de H ₂ em reatores anaeróbios em batelada.	25
Tabela 4: Principais vantagens da análise de regressão linear e não linear na digestão anaeróbia.....	30
Tabela 5: Variáveis composicionais, operacionais e respostas.	36
Tabela 6: Características gerais dos estudos.	36
Tabela 7: Testes Estatísticos.....	40
Tabela 8: Estatísticas descritivas (Modo Batelada).	41
Tabela 9: Matriz de Correlação de Spearman (Modo Batelada).....	43
Tabela 10: Matriz com p-valor dos coeficientes (Modo Batelada).	44
Tabela 11: Autovalores e Variância compartilhada pelas variáveis (Modo Batelada).....	47
Tabela 12: Score Fatoriais (Modo Batelada).	48
Tabela 13: Cargas Fatoriais (Modo Batelada).	49
Tabela 14: Comunalidades (Modo Batelada).....	50
Tabela 15: Estatísticas Univariadas (Modo Batelada).	52
Tabela 16: Coeficientes do Modelo Linear (Modo Batelada).	54
Tabela 17: Estatísticas resumidas do Modelo Linear (Modo Batelada).	55
Tabela 18: Coeficientes do Modelo Linear após Stepwise (Modo Batelada). .	55
Tabela 19: Estatísticas resumidas do Modelo Linear após Stepwise (Modo Batelada).....	56
Tabela 20: Teste de Shapiro-Francia Modelo Linear (Modo Batelada).	57
Tabela 21: Teste de Breusch-Pagan Modelo Linear (Modo Batelada).	57
Tabela 22: Diagnóstico de multicolinearidade Modelo Linear (Modo Batelada).	58
Tabela 23: Coeficientes do Modelo Não Linear (Modo Batelada).....	59
Tabela 24: Tabela 24: Estatísticas resumidas do Modelo Não Linear (Modo Batelada).....	59

Tabela 25: Coeficientes do Modelo Não Linear após Stepwise (Modo Batelada).	60
Tabela 26: Estatísticas resumidas do Modelo Não Linear após Stepwise (Modo Batelada).	60
Tabela 27: Teste de Shapiro-Francia Modelo Não Linear após Stepwise (Modo Batelada).	61
Tabela 28: Teste de Breusch-Pagan Modelo Não Linear após Stepwise (Modo Batelada).	61
Tabela 29: Diagnóstico de Multicolinearidade Modelo Não Linear após Stepwise (Modo Batelada).	62
Tabela 30: Coeficientes do Modelo com Interações (Modo Batelada).	65
Tabela 31: Estatísticas resumidas do Modelo com Interações (Modo Batelada).	66
Tabela 32: Coeficientes do Modelo com Interações após Stepwise (Modo Batelada).	67
Tabela 33: Estatísticas resumidas do Modelo com Interações Após Stepwise (Modo Batelada).	67
Tabela 34: Teste de Shapiro-Francia Modelo com Interações após Stepwise (Modo Batelada).	69
Tabela 35: Teste de Breusch-Pagan Modelo com Interações após Stepwise (Modo Batelada).	69
Tabela 36: Diagnóstico de Multicolinearidade Modelo com Interações após Stepwise (Modo Batelada).	70
Tabela 37: Estatísticas descritivas (Modo Contínuo).	72
Tabela 38: Matriz de correlações (Modo Contínuo).	73
Tabela 39: Matriz com p-valor dos coeficientes (Modo Contínuo).	74
Tabela 40: Autovalores e Variância compartilhada pelas variáveis (Modo Contínuo).	78
Tabela 41: Score Fatoriais (Modo Contínuo).	79
Tabela 42: Cargas Fatoriais (Modo Contínuo).	80
Tabela 43: Comunalidades (Modo Contínuo).	82
Tabela 44: Estatísticas Univariadas (Modo Contínuo).	86
Tabela 45: Coeficientes do Modelo Linear (Modo Contínuo).	88
Tabela 46: Estatísticas resumidas do Modelo Linear (Modo Contínuo).	88

Tabela 47: Coeficientes do Modelo Linear após Stepwise (Modo Contínuo)..	89
Tabela 48: Estatísticas resumidas do Modelo Linear após stepwise (Modo Contínuo).....	89
Tabela 49: Teste Shapiro-Francia do modelo Linear após stepwise (Modo Contínuo).....	90
Tabela 50: Teste Breusch-Pagan do modelo Linear após stepwise (Modo Contínuo).....	91
Tabela 51: Diagnóstico de Multicolinearidade do modelo Linear após stepwise (Modo Contínuo).	91
Tabela 52: Coeficientes do Modelo Não Linear (Modo Contínuo).	92
Tabela 53: Estatísticas resumidas do Modelo Não Linear (Modo Contínuo). .	92
Tabela 54: Coeficientes do Modelo Não Linear após Stpwise (Modo Contínuo).	93
Tabela 55: Estatísticas resumidas do Modelo Não Linear após Stewise (Modo Contínuo).....	93
Tabela 56: Teste Shapiro-Francia do Modelo Não Linear após Stepwise (Modo Contínuo).....	94
Tabela 57: Teste Breusch-Pagan do Modelo Não Linear após Stepwise (Modo Contínuo).....	94
Tabela 58: Diagnóstico de Multicolinearidade do Modelo Não Linear após Stepwise (Modo Contínuo).	95
Tabela 59: Coeficientes do Modelo com Interações (Modo Contínuo).	97
Tabela 60: Estatísticas resumidas do Modelo com Interações (Modo Contínuo).	98
Tabela 61: Coeficientes do Modelo com Interações após Stepwise (Modo Contínuo).....	99
Tabela 62: Estatísticas resumidas do Modelo com Interações após Stepwise (Modo Contínuo).	99
Tabela 63: Teste Shapiro-Francia do Modelo com Interações após Stepwise (Modo Contínuo).	101
Tabela 64: Teste Breusch-Pagan do Modelo com Interações após Stepwise (Modo Contínuo).	101
Tabela 65: Diagnóstico de Multicolinearidade do Modelo com Interações após Stepwise (Modo Contínuo).	102

Tabela 66: Resumo do Modelos.	106
------------------------------------	-----

LISTA DE ABREVIATURAS E SIGLAS

CHT - Carboidratos totais

DQO - Demanda química de oxigênio

ETE - Estação de tratamento de esgoto

HY - Produção acumulada de hidrogênio (para reatores em batelada)

HY - Rendimento de hidrogênio (para reatores operados em modo contínuo)

LCA - Laboratório de Controle Ambiental

LSA - Laboratório de Saneamento Ambiental

N - Nitrogênio

pH - Potencial hidrogeniônico

P - Fósforo

SI - Razão substrato-inóculo

ST - Sólidos totais

STF - Sólidos totais fixos

STV - Sólidos totais voláteis

SVI - Concentração de sólidos voláteis do inóculo

t - Tempo

TCO - Taxa de carga orgânica

TDH - Tempo de detenção hidráulica

UASB - Reator anaeróbio de fluxo ascendente de manta de lodo

UFAL - Universidade Federal de Alagoas

VI - Volume de inóculo

SUMÁRIO

1. INTRODUÇÃO	16
2. OBJETIVOS E METAS	18
2.1. Objetivo geral	18
2.2. Objetivos específicos.....	18
3. REVISÃO BIBLIOGRÁFICA	19
3.1. Vetor Energético: Hidrogênio.....	19
3.2. Produção biológica de hidrogênio.....	20
3.3. Aspectos gerais do processo de fermentação anaeróbia	21
3.4. Fatores que influenciam o processo fermentativo anaeróbio.....	22
3.4.1. pH	22
3.4.2. Temperatura	23
3.4.3. Tipo de Inóculo	24
3.4.4. Substratos.....	26
3.4.5. Modo Operacional do Reator	26
3.5. Modelagem de Reatores Anaeróbios.....	28
3.5.1. Análise de Componentes Principais (ACP) ou Principal Component Analysis (PCA) em Reatores Anaeróbios.....	30
3.5.2. Análise de Regressão Linear e Não Linear na Modelagem da Digestão Anaeróbia.....	32
4. METODOLOGIA.....	35
4.1. Determinação das variáveis e Agrupamento de dados	35
4.2. Aspectos gerais dos estudos.....	36
4.3.1. Análise de Componentes Principais (PCA).....	38
4.3.2. Análise de Regressão Múltipla Linear e Não Linear	39
5. RESULTADOS E DISCUSSÕES	41
5.1. Análise Fatorial por Componentes Principais (Reatores operados em batelada).....	41
5.1.1. Estatísticas descritivas.....	41
5.1.2. Coeficientes de correlação de Spearman para cada par de variáveis ..	42
5.1.3. Elaboração a Análise Fatorial Por Componentes Principais	46
5.2. Análise de Regressão para modo Batelada	51
5.2.1. Estatísticas Univariadas.....	51
5.2.2. Regressão Linear Múltipla	54
5.2.3. Modelo de Regressão Não Linear Usando Transformação de BOX-COX	59
5.2.4. Análise De Regressão Múltipla Com Interações De Segunda Ordem ..	64

5.3. Análise Fatorial por Componentes Principais para variáveis operadas com reator em modo contínuo	72
5.3.1. Estatísticas descritivas.....	72
5.3.2. Coeficientes de correlação de Spearman para cada par de variáveis ..	73
5.3.3. Elaboração da Análise Fatorial Por Componentes Principais	77
5.4. Análise de Regressão Múltipla (Variáveis operadas em reatores contínuos)	86
5.4.1. Estatísticas Univariadas.....	86
5.4.2. Estudo das Correlações.....	87
5.4.3. Regressão Linear Múltipla	88
5.4.4. Modelo Regressão Não Linear Usando Transformação de BOX-COX.	92
5.4.5. Análise De Regressão Múltipla Com Interações De Segunda Ordem ..	97
5.5 Síntese dos Resultados	105
6. CONCLUSÕES	107
REFERÊNCIAS	109

1. INTRODUÇÃO

Os combustíveis fósseis, além de serem fontes de energia não renováveis, são responsáveis por diversos impactos ambientais, como o crescente aquecimento global (Veziroglu, Sahin, 2008). Nesse contexto, na exploração de fontes alternativas e mais sustentáveis de energia, o hidrogênio tem ganhado cada vez mais destaque (Benvindo, 2023).

O hidrogênio é considerado um vetor energético bastante viável, visto que é o elemento mais abundante do planeta, sua combustão direta gera uma quantidade significativa de energia, é uma fonte de energia limpa, e é considerado o combustível mais leve e o mais rico em energia por unidade de massa (Madeira, 2018).

No que se refere aos processos de obtenção de hidrogênio, a produção biológica por meio de fermentação anaeróbia é uma opção promissora, pois os processos fermentativos apresentam a grande vantagem de possibilitar a utilização de biomassa e de inúmeros materiais residuais como substrato (Coêlho, 2013).

Diante do contexto atual, a problemática acerca dos resíduos industriais está diretamente ligada ao atraso do progresso econômico, bem como à falta de soluções sanitárias e ambientalmente corretas em relação à disposição final ou reaproveitamento desses resíduos (Farage *et al.*, 2013). E ressaltando o grande potencial desses materiais ricos em carboidratos (o que favorece o processo de bioconversão) como fonte para geração de biocombustíveis, é essencial que cada vez mais se invista no desenvolvimento de novos estudos acerca dessa questão. Tais iniciativas podem promover benefícios ambientais ao minimizar os potenciais impactos de seu descarte sem pré-tratamento (Cappelletti, 2009), além da geração de biocombustíveis.

Muitos fatores podem influenciar o processo de produção do bio-hidrogênio, como a quantidade de substrato e sua composição química, o tipo e quantidade de inóculo, a temperatura, o pH, a relação carbono/nitrogênio, os micronutrientes (Braguglia *et al.*, 2018), bem como os pré-tratamentos físico-químicos. O gerenciamento desses elementos tem sido empregado para aprimorar as características morfológicas e estruturais de diferentes biomassas (Dutra *et al.*, 2020; Dutra, *et al.*, 2022; Poddar *et al.*, 2021).

Atualmente inúmeros estudos têm avaliado os parâmetros para otimização da produção de hidrogênio a partir da fermentação utilizando vários resíduos industriais

diferentes no processo, demonstrando a importância do tipo de substrato utilizado no rendimento final da produção de H₂ (Lamaison, 2009). Porém, a literatura é escassa em termos de comparações em função das suas variáveis composicionais, operacionais e respostas dos processos avaliados.

Assim sendo, o presente estudo apresenta uma análise de dados para que possa ser realizada uma análise crítica acerca dos trabalhos realizados entre os anos de 2013 a 2023 de estudos realizados pelo grupo de pesquisa do Laboratório de Controle Ambiental, do Centro de Tecnologia da Universidade Federal de Alagoas. As pesquisas, orientadas pelo Prof. Dr. Eduardo Lucena Cavalcante de Amorim, exploram a digestão anaeróbia de diversos resíduos da região semiárida de Alagoas, utilizando o mesmo tipo de inóculo, tendo como variáveis resposta o nível de rendimento e a produção acumulada de hidrogênio. Deste modo, o foco é evidenciar uma tecnologia sustentável como alternativa aos combustíveis fósseis, usando esses resíduos agroindustriais na produção de bio-hidrogênio.

Poucos estudos, como o de Dutra *et al.* (2023), que realiza uma revisão sistemática e meta-análise sobre a digestão anaeróbia utilizando resíduos de cacau, abordam de forma integrada as variáveis operacionais e composicionais, avaliando como essas influenciam a produção de energia. Nesse contexto, o presente estudo busca preencher essa lacuna, correlacionando as variáveis operacionais e composicionais com a produtividade de hidrogênio, o que constitui a principal contribuição científica e tecnológica deste trabalho. Técnicas quantitativas, incluindo métodos estatísticos de análise multivariadas como a Análise de Componentes Principais, Análise de Regressão Múltipla e meta-análise, também foram utilizadas.

Resultados preliminares indicam que a interação entre as variáveis operacionais e composicionais é crucial para otimizar a produção de hidrogênio. O modelo de predição com interações de segunda ordem apresentou-se como o mais eficaz para estimar a produtividade de bio-hidrogênio, revelando a importância de ajustes finos nas condições operacionais. Este trabalho, portanto, não só contribui para a compreensão dos fatores que impactam a produção de hidrogênio, mas também oferece uma base sólida para otimizações futuras em processos de digestão anaeróbia, ampliando a aplicação de resíduos agroindustriais como fonte sustentável de energia.

2. OBJETIVOS E METAS

2.1. Objetivo geral

Avaliar o potencial de resíduos agroindustriais para a produção de hidrogênio por meio de digestão anaeróbia, utilizando métodos de análise multivariada para investigar o impacto das variáveis operacionais e composicionais nos processos.

2.2. Objetivos específicos

- Analisar os dados existentes sobre o processo de digestão anaeróbia para produção de hidrogênio a partir de estudos anteriores da equipe de pesquisa do LCA, com foco nos reatores contínuos e em batelada;
- Identificar e avaliar as principais correlações entre as variáveis operacionais que influenciam a produção de hidrogênio;
- Avaliar e comparar a eficácia de modelos de regressão linear e não linear na predição da produção de hidrogênio, visando otimizar os processos de digestão anaeróbia.

3. REVISÃO BIBLIOGRÁFICA

3.1. Vetor Energético: Hidrogênio

Segundo EIA (2016), o consumo de energia mundial, no contexto de fontes energéticas, irá aumentar em 48% até o ano de 2040, e dessa porcentagem, os combustíveis fósseis ainda continuarão representando 80% dessa demanda. Contudo, sabe-se da necessidade da procura por fontes de energia renováveis e sustentáveis devido à depleção das reservas fósseis associada aos danos ambientais provocados pela sua utilização.

Na busca por fontes alternativas e sustentáveis de energia, o hidrogênio tem sido amplamente reconhecido como um vetor energético promissor. Sua abundância no planeta, aliada à capacidade de combustão limpa e altamente energética, torna-o uma opção atrativa para reduzir emissões de carbono e mitigar o impacto ambiental associado aos combustíveis fósseis (Madeira, 2018; Veziroglu, Sahin, 2008). Além de ser o elemento mais leve e com a maior densidade energética por unidade de massa, o hidrogênio se destaca por suas aplicações em tecnologias limpas, o que o posiciona como um combustível ideal em estratégias de transição energética sustentável (Braguglia *et al.*, 2018).

Seu potencial energético é bastante elevado (122 kJ g^{-1}), sendo maior que o do etanol ($26,5 \text{ kJ g}^{-1}$) e do metano ($50,1 \text{ kJ g}^{-1}$). Além disso, possui uma densidade energética cerca de 2,75 vezes superior à do petróleo (Botta, 2012). Outro aspecto positivo do hidrogênio é a alta flexibilidade tecnológica em relação aos seus processos de produção a partir de matérias-primas renováveis ou não renováveis (Page; Krumdieck, 2009; Sá *et al.*, 2014).

Na área de desenvolvimento de tecnologias para a produção de hidrogênio, entra a produção biológica desse elemento como um campo promissor e em destaque. Quando produzido com a finalidade de ser usado como biocombustível a partir da fração biodegradável de resíduos, ou a partir de uma biomassa, é chamado também de bio-hidrogênio. Sua produção, quando associada ao tratamento de resíduos orgânicos, engloba também a minimização e tratamento desse material.

3.2. Produção biológica de hidrogênio

Segundo Das e Veziroğlu (2001) pode-se obter a produção biológica de hidrogênio (bio-hidrogênio) através de diferentes processos: fermentação anaeróbia ou fermentação escura de compostos orgânicos por meio de bactérias fermentativas e sistemas híbridos utilizando bactérias fotossintéticas e fermentativas, foto-fermentação de compostos orgânicos por bactérias fotossintetizantes, biofotólise da água, realizada por algas verdes (direta) e cianobactérias (indireta).

O processo de fermentação anaeróbia ou fermentação escura de compostos orgânicos por meio de bactérias fermentativas e sistemas híbridos utilizando bactérias fotossintéticas e fermentativas envolve de modo geral bactérias que transformam carboidratos em hidrogênio (Argun; Kargi, 2010). A foto-fermentação utiliza bactérias púrpuras não sulfurosas que usam energia luminosa para converter ácidos orgânicos em hidrogênio e gás carbônico (Sá *et al.*, 2014). Na biofotólise direta, as algas verdes usam a energia da luz para quebrar as moléculas de água em hidrogênio e oxigênio (O_2) sob condições anaeróbias. Na biofotólise indireta, as cianobactérias usam a energia armazenada nos carboidratos produzidos durante a fotossíntese para formar H_2 a partir da água (Brentner; Peccia; Zimmerman, 2010).

A Tabela 1 apresenta as principais vantagens e desvantagens desses processos biológicos:

Tabela 1: Vantagens, desvantagens e principais micro-organismos estudados para produção de hidrogênio biológico

Processos biológicos	Vantagens	Desvantagens	Micro-organismos
Biofotólise direta	Produção de H_2 a partir da água; Não necessita de ATP.	Necessita de iluminação constante; Inibição das hidrogenases pelo O_2 ; Biorreatores modificados para maximizar a utilização da energia solar	<i>Chlamydomonas reinhardtii</i> <i>Platymonas subcordiformis</i>
Biofotólise indireta	Produção de H_2 a partir da água; Habilidade de fixação de N_2 e produção de H_2 pelas nitrogenases; Separação espacial do O_2 e H_2 nas células diferenciadas (heterocistos); Separação temporal (dia/noite) em células não-heterocísticas.	Necessita de iluminação constante; Biorreatores modificados para maximizar a utilização da energia solar; Necessidade de ATP pelas nitrogenases; Síntese e manutenção dos heterocistos; CO_2 presente no gás.	<i>Plectonema boryanum</i> <i>Anabaena siamensis</i> <i>Anabaena variabilis</i> <i>Synechocystis sp.</i> <i>Cyanothece sp.</i> <i>Nostoc sp.</i>

Foto-fermentação	Utilização de diferentes resíduos e efluentes como substrato; Utilização de amplo espectro de luz pelas bactérias fotossintéticas.	Necessita de iluminação; Biorreatores modificados para maximizar a utilização da energia solar; CO ₂ presente no gás.	<i>Rhodopseudomonas palustris</i> <i>Rhodobacter sp.</i> <i>Rhodobacter sphaeroides</i> <i>Rhodobacter capsulatus</i>
Fermentação anaeróbia	Ampla variedade de fontes de carbono como substrato; Não necessita de iluminação; Produção de metabólitos intermediários de grande valor agregado; Não há limitações com relação ao O ₂ .	Tratamento do efluente após fermentação; CO ₂ presente no gás.	<i>Clostridium sp.</i> <i>Clostridium butyricum</i> <i>Clostridium beijerinckii</i> <i>Citrobacter freundii</i> <i>Enterobacter cloacae</i> <i>Enterobacter aerogenes</i> <i>Escherichia coli</i> <i>Klebsiella pneumoniae</i>

Fonte: Adaptado de Sá, Cammarota e Ferreirão-Leitão (2014).

3.3. Aspectos gerais do processo de fermentação anaeróbia

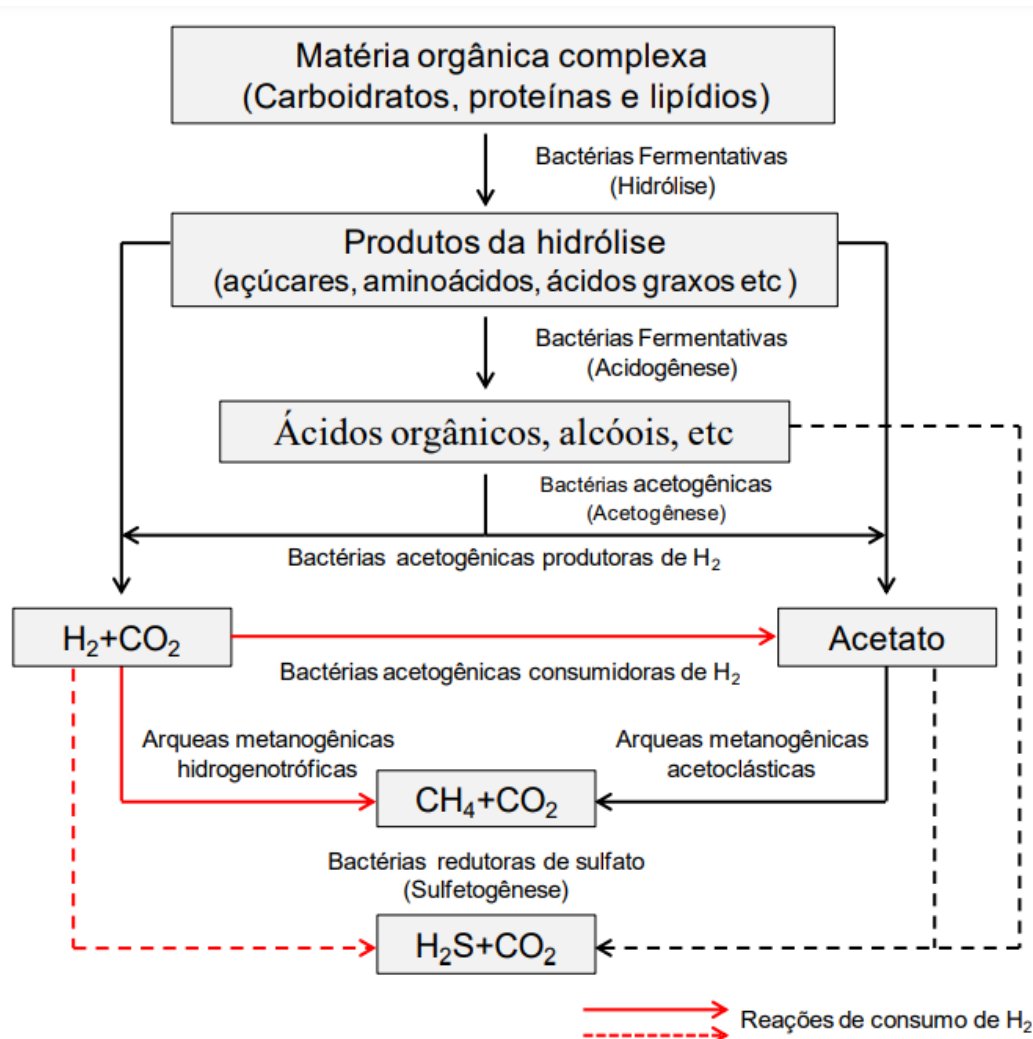
A fermentação anaeróbia, como já citado anteriormente, é um processo biológico onde um conjunto de diferentes tipos de microrganismos possibilita a transformação de compostos orgânicos complexos (carboidratos, lipídios e proteínas) em produtos mais simples, tais como ácidos orgânicos voláteis, álcoois, H₂, CO₂ e CH₄ (Sá; Cammarota; Ferreira-Leitão; 2014).

É considerado um processo complexo que envolve quatro fases principais: hidrólise, acidogênese, acetogênese e metanogênese (Figura 1). Na hidrólise (primeira fase do processo), encontram-se bactérias hidrolíticas que são capazes de produzir enzimas extracelulares que permitem que materiais complexos sejam decompostos em materiais mais simples (Chernicharo, 2007).

Posteriormente, na fase acidogênica, as bactérias fermentativas irão converter os compostos gerados na fase de hidrólise em ácidos orgânicos e álcoois. Vale ressaltar também que essas bactérias acetogênicas são responsáveis pela oxidação dos produtos gerados nessa segunda fase, onde os ácidos são convertidos em hidrogênio, gás carbônico e acetato (Chernicharo, 2007).

Na terceira fase, as bactérias acetogênicas são responsáveis pela oxidação dos produtos gerados na segunda fase em substratos adequados para as arqueias metanogênicas. Nessa fase do processo, o hidrogênio também pode ser transformado em ácido acético pelas bactérias homoacetogênicas. E, por fim, na última fase, as arqueias metanogênicas convertem o hidrogênio e o ácido acético em CH₄ e CO₂ (Sá; Cammarota; Ferreira-Leitão, 2014).

Figura 1: Representação esquemática das etapas e microrganismos envolvidos no processo de digestão anaeróbia, bem como as reações de consumo de hidrogênio.



Fonte: Silva (2017).

3.4. Fatores que influenciam o processo fermentativo anaeróbio

A produção do bio-hidrogênio é afetada por uma série de parâmetros e condições que podem influenciar diretamente o metabolismo microbiano e o rendimento do processo. Entre os parâmetros mais influentes estão o pH inicial, temperatura, tipo de inóculo, substrato e modo operacional do reator (Ghimire *et al.*, 2015; Intanoo *et al.*, 2014; Soares, 2019).

3.4.1. pH

Diversos estudos apontam que o pH ótimo, em termos de produção de hidrogênio, está entre 5 e 7 (Khanal, 2004; Ginkel, 2001 apud Silva, 2017). Esse pH favorece a atividade de enzimas envolvidas na produção de hidrogênio e, ao mesmo

tempo, permite inibir microrganismos que consomem o H_2 presente no meio de fermentação (Whang, 2005; Valdez-Vazquez; Poggi-Varaldo, 2009).

Nualsri *et al.* (2016) ressaltam que bactérias acidogênicas preferem a faixa de pH de 5,5 a 6,5 e que valores muito inferiores ou superiores a esta faixa minimizarão a produção de hidrogênio. Enquanto, de acordo com Ohimain e Izah (2017), o pH neutro promoveria a produção de metano durante a produção de biogás, e a maioria das bactérias anaeróbias, especialmente as metanogênicas, aumentam a produção de metano em pH 6,5 a 7,5, reduzindo assim a produção do bio-hidrogênio. Portanto, para maior produção de hidrogênio e com o intuito de evitar a geração de metano, uma faixa de pH de 5,5 a 6,5 é considerada ideal. Vale ressaltar que, segundo estudos anteriores, um ambiente muito ácido, com pH abaixo de 4,5, inibe a atividade das enzimas envolvidas na hidrogenação (hidrogenase) (Peixoto, 2011; Bharathiraja *et al.*, 2016).

Na Tabela 2, são apresentadas as faixas de pH aplicadas nos experimentos para produção de hidrogênio em reatores anaeróbios em batelada.

Tabela 2: Comparação entre as faixas de pH aplicadas e o pH ótimo obtido na produção de H_2 em reatores anaeróbios em batelada.

Faixa pH	pH Ótimo	Controle de pH	Referência
4,5; 5,5 e 6,5	5,5	Sim	Shin, Youn e Kim (2004)
5,5; 6,0; 6,5 e 7,0	7,0	Sim	Lee <i>et al.</i> (2008)
5,0; 6,0; 7,0; 8,0 e 9,0	7,5	NR	Ismael <i>et al.</i> (2009)
1,0 e 13,0; 7,0	NR	Sim	Kim, Kim e Shin (2009)
5,0; 6,0; 7,0 e 8,0	7,0	NR	Nazlina <i>et al.</i> (2009)
5,0; 5,5 e 6,0	5,5	Sim	Yasin <i>et al.</i> (2011)
5,5; 6,0 e 7,0	7,0	Sim	Ramos <i>et al.</i> (2012)

Fonte – Adaptado de Lima, 2019.

3.4.2. Temperatura

A temperatura é um dos fatores de principal influência na produção de H_2 . As reações de fermentação podem ser realizadas e divididas em quatro faixas principais de temperatura: mesofílica (25-40 °C), termofílica (40-65 °C), termofílica extrema (65-80 °C) e hipertermofílica (> 80 °C) (Levin, 2004).

O processo de produção de hidrogênio a partir de resíduos orgânicos é realizado em condições de temperatura ambiente (mesofílicas ou termofílicas). Bactérias acidogênicas tendem a ter maiores possibilidades de crescimento sendo

conduzidas nessas faixas de temperatura; porém, algumas espécies também podem sobreviver a baixas temperaturas (Karadag *et al.*, 2014). A temperatura ótima para cada processo depende do tipo e origem do inóculo, da quantidade de compostos biodegradáveis e das condições operacionais do sistema (Silva, 2017).

No entanto, vale ressaltar que a maioria dos estudos dedicados à investigação em laboratório sobre a produção de hidrogênio com culturas mistas foi realizada em condições mesofílicas (25-40 °C). Levin, Pitt e Love (2004) relataram que em condições mesofílicas houve uma maior variedade de microrganismos que produzem hidrogênio. Essa faixa de temperatura é mais fácil de ser controlada em escala laboratorial, além da vantagem de se ter menos gastos de energia, permitindo que atinjam bons resultados (Kothari, 2014; Mao *et al.*, 2015; Yasin *et al.*, 2013).

3.4.3. Tipo de Inóculo

Em reatores para a produção do bio-hidrogênio, uma grande variedade de microrganismos anaeróbios é capaz de produzir H_2 via fermentação anaeróbia, sob a forma de culturas puras de bactérias produtoras de hidrogênio, combinação de culturas puras (coculturas), consórcio microbiano (culturas mistas), microrganismos geneticamente modificados e culturas puras. No entanto, as culturas puras e mistas são as mais estudadas (Wang; Wan, 2011; Nunes, 2015).

As culturas puras permitem a manipulação do metabolismo microbiano, possibilitando modificações nas condições de crescimento e aumento na produtividade do H_2 (Ntaikou *et al.*, 2010; Tonello, 2017). Contudo, nos processos de produção via fermentação do bio-hidrogênio, as culturas mistas são mais frequentes, devido às vantagens apresentadas em relação a questões operacionais e econômicas, ao passo que também permitem a possibilidade de se usar uma vasta diversidade de substrato (Wang; Wan, 2009).

Em contrapartida, culturas mistas também estão associadas à existência de bactérias metanogênicas que são capazes de consumir o H_2 gerado pelas bactérias fermentativas acidogênicas (Show *et al.*, 2012; Tonello, 2017). Em função disso, e visando a mitigação dessa consequência, faz-se necessário um pré-tratamento do inóculo com a intenção de inibir ou eliminar esses microrganismos consumidores de H_2 .

Segundo Cisneros-Pérez *et al.* (2017), os procedimentos mais comuns utilizados no pré-tratamento do inóculo são os tratamentos térmicos, químicos e

elétricos. Nesse contexto, o uso do método térmico tem funcionado de maneira eficiente na inativação de bactérias metanogênicas (Barros *et al.*, 2010; De Gioannis *et al.*, 2013; Santos *et al.*, 2014). Além disso, o tratamento térmico também tem sido apontado como o mais eficiente em relação ao maior e melhor rendimento de produtividade de H₂ (Shida *et al.*, 2009; Tonello, 2017).

Na Tabela 3 são apresentados os pré-tratamentos comumente aplicados no inóculo para produção de hidrogênio em reatores anaeróbios em batelada.

Tabela 3: Comparação entre pré-tratamentos do inóculo aplicados na produção de H₂ em reatores anaeróbios em batelada.

Inóculo [Origem]	Pré-tratamento	Referência
Lodo de esgoto [Digestor anaeróbico]	Térmico (10 min a 90°C)	Kim, Han e Shin (2004)
Cultura acidogênica [CSTR]	Lavagem com meio anaeróbio	Shin, Youn e Kim (2004)
Gramma biodigerida Anaerobicamente	Térmico (2 a 8 h e 60 a 120°C)	Lay <i>et al.</i> (2005)
Lodo de esgoto [Digestor anaeróbico]	Térmico (15 min a 80°C)	Ming <i>et al.</i> (2008)
Lodo anaeróbio [Tanque de decantação]	Térmico (20 min a 80°C)	Ismail <i>et al.</i> (2009)
Inóculo aclimatado (50 mL de lodo anaeróbio [Digestor anaeróbio] tratado termicamente + 50 mL de meio Clostridial)	Térmico (30 min a 80°C)	Nazlina <i>et al.</i> (2009)
Lama de rio	Térmico (Fervido por 15 min)	Wang <i>et al.</i> (2010)
Lodo anaeróbio	Térmico (20 min a 90°C/2 h a 550°C)	Kim <i>et al.</i> (2011)
Inóculo aclimatado (50% (m/v) de lodo anaeróbio tratado termicamente + 50% (m/v) de substrato)	Térmico (30 min a 80°C)	Yasin <i>et al.</i> (2011)
Lodo granular anaeróbio	Térmico (24 h a 104°C)	Ramos <i>et al.</i> (2012)
Lodo anaeróbio [CSTR]	Térmico (10 min a 105°C)	Abubackar <i>et al.</i> (2019)
Lodo anaeróbio [Digestor anaeróbico]	2-bromoetano sulfônico, térmica,	Basak <i>et al.</i> (2019)
Lodo anaeróbio [CSTR]	ácido, básico, UV e ultrassom Térmico (10 min a 105°C)	Keshin <i>et al.</i> (2019)

Lodo anaeróbico [Planta de tratamento de esgoto]	Térmico (Fervido por 1 h)	Turhal, Turanbaev e Argun (2019)
--	---------------------------	-------------------------------------

Fonte – Adaptado de Lima, 2019.

3.4.4. Substratos

O tipo de substrato influencia fortemente a viabilidade técnica e econômica da produção de hidrogênio por fermentação anaeróbia. Nesse contexto, uma diversificada quantidade de materiais ricos em carboidratos, proteínas e/ou lipídios pode ser utilizada como fonte de substratos. Contudo, entre esses fatores, os carboidratos possuem maior influência no processo fermentativo.

Segundo Lay (2003), a produção de hidrogênio a partir de resíduos com alto teor de carboidratos possui um potencial de produção cerca de 20 vezes maior do que os resíduos ricos em proteínas e/ou gorduras. Esse fator está associado à produção de amônia devido ao acúmulo de ácidos graxos de cadeia longa pela degradação de lipídios, bem como à degradação de proteínas. Ambos são capazes de inibir as bactérias acidogênicas quando presentes em altas concentrações (Dong *et al.*, 2009).

Quanto às características de um potencial substrato, Lamaison faz a seguinte afirmação:

Existem alguns critérios que devem ser obedecidos para que o substrato seja considerado adequado para ser utilizado na produção de hidrogênio. São eles: ser rico em carboidratos, ser produzido a partir de recursos renováveis, estar em concentração suficiente para que a fermentação e a recuperação da energia sejam energeticamente favoráveis e que, no caso de algum tratamento prévio ser necessário, este seja de baixo custo (Lamaison, 2009).

Assim, vários materiais residuais podem e têm sido utilizados como substrato para as bactérias fermentativas na produção de hidrogênio.

3.4.5. Modo Operacional do Reator

O reator é considerado um dos parâmetros fundamentais no processo de produção biológica do bio-hidrogênio, de modo que a configuração afeta diretamente as condições do microambiente interno do reator. Estudos indicam que, na produção de H₂ por fermentação anaeróbia, reatores operados em batelada são particularmente adequados em escala laboratorial, uma vez que oferecem simplicidade, flexibilidade

e facilidade de uso (Chen *et al.*, 2008; Hawkes *et al.*, 2007; Lin; Lay, 2004). Esses reatores permitem um controle mais preciso das condições operacionais e facilitam o estudo de parâmetros específicos, favorecendo assim o desenvolvimento de metodologias para otimização do processo.

Contudo, em um contexto industrial, operações em larga escala demandam um processo de produção contínua, por motivos de praticidade, armazenamento e questões econômicas (Guo *et al.*, 2010). Na maioria dos trabalhos utilizados com fins de investigação laboratorial, utilizam-se frascos de pequenos volumes, entre 100 mL e 500 mL, e fermentadores entre 2 mL e 10 mL com agitação (De Gioannis *et al.*, 2013).

No processo contínuo em escala laboratorial, o reator de tanque agitado contínuo (continuous stirred tank reactor – CSTR) é amplamente utilizado devido à sua capacidade de manter uma mistura homogênea e condições constantes de operação (Show *et al.*, 2011; Lee *et al.*, 2010). Outros tipos de reatores também têm sido aplicados, como o reator anaeróbio de fluxo ascendente e manta de lodo (upflow anaerobic sludge blanket – UASB), biorreator de leito empacotado, reator de leito fixo, biorreator de leito fixo e fluxo ascendente, reator batelada anaeróbio em série e o reator de leito fluidizado anaeróbio (Gupta *et al.*, 2016; Sarma *et al.*, 2013).

Entre essas configurações, destaca-se também o reator de alta taxa de fluxo ascendente (RALF – rapid anaerobic liquid flow reactor), que tem sido explorado por sua eficiência na remoção de compostos orgânicos e produção de biocombustíveis, incluindo o hidrogênio. Os reatores RALF operam em condições de alta taxa de fluxo e volume de reciclagem, permitindo uma mistura mais eficiente e reduzindo o tempo de retenção hidráulica, o que é vantajoso para aplicações em larga escala e operações contínuas (Kim *et al.*, 2011; Luo *et al.*, 2019). Essa configuração proporciona um ambiente adequado para o crescimento de culturas microbianas eficientes, aumentando o rendimento do hidrogênio e reduzindo o acúmulo de subprodutos inibidores no processo fermentativo.

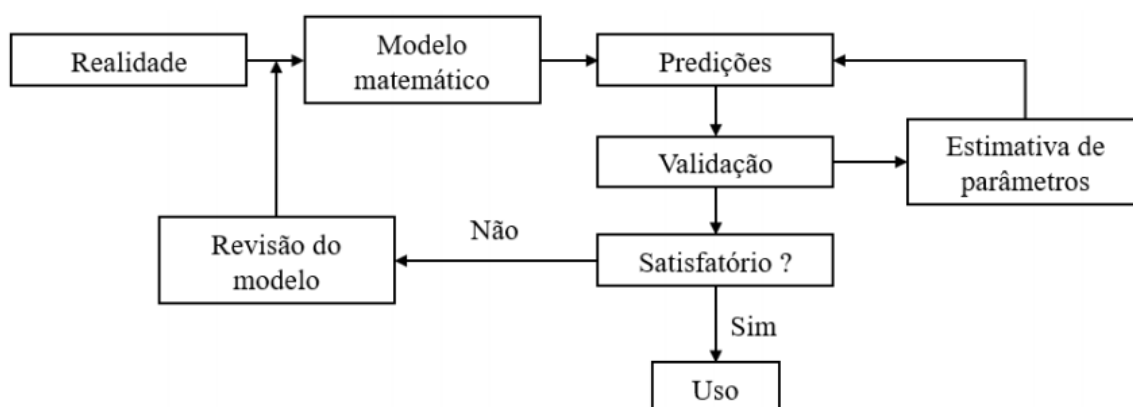
Em estudos de produção de hidrogênio, o processo em batelada demonstrou maiores rendimentos de H_2 em comparação ao modo contínuo. Esse fenômeno é frequentemente atribuído ao microambiente diferenciado no reator em batelada, que oferece condições propícias para a atividade microbiana eficiente em termos de produção de hidrogênio (Lay *et al.*, 1999; Das; Veziroglu, 2001).

3.5. Modelagem de Reatores Anaeróbios

A criação de modelos matemáticos baseados em resultados de experimentos em laboratório é uma maneira de coletar informações sobre o funcionamento de reatores variados, definir seus parâmetros e avaliar seu rendimento, com um custo reduzido e maior flexibilidade em comparação à construção de protótipos (Neckel, 2018).

Os processos fermentativos geralmente envolvem o controle de variáveis ambientais, reações simultâneas e uma maneira de estruturar esses processos de forma a respeitar o regime de escoamento. Esses acontecimentos podem ser retratados por modelos matemáticos, que, dentro de suas restrições, estruturam equações matemáticas a partir de simplificações de equações constitutivas, ponderações sobre os princípios de conservação e erros de medida, além da habilidade de calcular (Mendes; Valdman; Souza; Júnior, 2011). As resoluções de sistemas de equações fornecem previsões sobre o modelo, que quando comparado aos dados experimentais, seja pela aproximação dos dados simulados aos experimentais ou através do emprego de algum protocolo estatístico, permitem validar ou não o modelo para uso direto (Nascimento, 2021).

Figura 2: Proposta de desenvolvimento de modelo matemático.



Fonte: adaptado de Mendes, Valdman e Souza Júnior (2011).

Em estações de tratamento de efluentes, os modelos podem ser empregados para aprimorar e testar estratégias de controle dos processos, com o objetivo de cumprir os padrões de qualidade a custos razoáveis (Turkdogan-Aydinol, 2011). Hill e colaboradores (1977), desenvolveram o primeiro modelo de simulação de sistemas anaeróbios, em um reator de mistura completa, utilizando duas populações

microbianas, fermentadores e metanogênicos para transformar resíduos de produção de frango. Os autores tentaram ajustar o modelo realizando os experimentos através de doze bateladas, à temperatura de 25°C, obtendo sucesso em apenas oito delas. Após, examinaram como o modelo reagiria se a temperatura fosse reduzida para 15°C, e tiveram sucesso em apenas três das doze simulações realizadas. Assim, consideraram adequado calibrar o modelo à temperatura de 25°C, e posteriormente procuraram validar o modelo contra um reator em escala completa. A validação foi concluída com êxito, apresentando resultados com uma margem de erro de apenas 10% para as variáveis de controle: biomassa, ácidos voláteis e compostos de nitrogenados.

As propostas mais significativas para validar modelos que abordam os eventos da digestão anaeróbia baseiam-se na comparação de dados obtidos em simulações com os dados obtidos em experimentos. Frequentemente, a confirmação de um modelo matemático se baseia em parâmetros que não foram obtidos experimentalmente, os parâmetros livres. Esses parâmetros podem ser constantes ou variar dentro de um intervalo definido (Nascimento, 2021).

O modelo ADM1 (Anaerobic Digestion Model No.1) é amplamente utilizado para representar a dinâmica bioquímica complexa do processo, integrando análises de sensibilidade para identificar variáveis influentes. Porém, embora o ADM1 seja um modelo abrangente e amplamente aceito para a digestão anaeróbia, a análise de regressão linear e não linear oferece várias vantagens em termos de simplicidade, flexibilidade e capacidade de interpretação. Esses modelos podem ser particularmente úteis em estudos onde a complexidade dos dados e a necessidade de modelos rápidos e eficientes são prioridades.

A adoção de métodos de regressão para facilitar a compreensão e a otimização dos processos de digestão anaeróbia, especialmente na produção de biogás e bio-hidrogênio a partir de resíduos orgânicos é amplamente apoiado pela literatura científica. A Tabela 4 apresenta as principais vantagens da análise de regressão linear e não linear na digestão anaeróbia.

Tabela 4: Principais vantagens da análise de regressão linear e não linear na digestão anaeróbia.

Vantagens	Descrição	Referência
Simplicidade e Interpretação	Modelo Intuitivo: A análise de regressão, especialmente a linear, é mais simples e intuitiva, permitindo uma interpretação mais direta dos resultados. Coeficientes de regressão podem ser facilmente interpretados como mudanças na produção de biogás em resposta a alterações nas variáveis independentes.	Kafle, G.K., & Kim, S.H. (2013).
Menor Demanda de Dados	Menos Parâmetros: Os modelos de regressão geralmente requerem menos dados em comparação com o ADM1, que é complexo e exige um grande número de parâmetros cinéticos e estequiométricos.	Ostermeyer, Y., <i>et al.</i> (2018).
Flexibilidade em Relação a Dados Não Lineares	Modelagem de Relações Complexas: Modelos de regressão não linear permitem a modelagem de relações complexas entre variáveis, frequentemente observadas em processos de digestão anaeróbia, onde a produção de biogás pode não seguir uma relação linear com as variáveis de entrada.	Sulaiman, SA, <i>et al.</i> (2021).
Facilidade de Validação e Ajuste	Ajuste Simples: Modelos de regressão podem ser facilmente ajustados e validados com dados experimentais, permitindo que pesquisadores explorem diferentes configurações operacionais sem a necessidade de uma estrutura complexa como a do ADM1.	Mansaray, KG e Kazi, S. (2020).
Utilização em Análises de Sensibilidade	Identificação de Variáveis Críticas: A análise de regressão permite realizar análises de sensibilidade que ajudam a identificar quais variáveis têm maior impacto na produção de biogás, facilitando a otimização do processo.	Cai, M., <i>et al.</i> (2018).

Fonte: Autor (2024)

3.5.1 Análise de Componentes Principais (ACP) ou Principal Component Analysis (PCA) em Reatores Anaeróbios

A Análise de Componentes principais é uma análise multivariada, um modelo fatorial, na qual os componentes se baseiam na variância total. O objetivo dessa técnica é realizar o agrupamento de variáveis que apresentem coeficientes de correlação elevados, a partir dos quais são realizadas combinações lineares para gerar fatores (novo conjunto de variáveis). A análise dos componentes principais é feita utilizando as unidades na diagonal da matriz de correlação, implicando computacionalmente que toda a variância é comum ou compartilhada (Araujo, Coelho, 2009).

Segundo Fávero e Belfiore (2017), a PCA é uma redução estrutural também conhecida como transformação de Karhunen-Loève. Esse tipo de análise é utilizado para variáveis métricas e não tem caráter preditivo, podendo ser aplicado de forma simplificada em linguagens de programação como o R. A PCA envolve, resumidamente, as seguintes etapas:

I) Verificação das correlações entre as variáveis por meio da correlação linear de Pearson, definindo-se uma matriz de correlações;

II) Verificação da adequação global da análise por meio do teste de esfericidade de Bartlett. Essa etapa consiste na comparação entre a matriz de correlações (ρ) e a matriz identidade (I) de mesma dimensão, de modo que os valores fora da diagonal principal de cada matriz devem ser estatisticamente diferentes de zero;

III) Determinação dos fatores por componentes principais a partir da determinação de autovalores e autovetores da matriz de correlações e cálculo dos scores fatoriais;

IV) Escolha da quantidade de fatores por meio do critério da raiz latente ou critério de Kaiser, considerando-se apenas os fatores correspondentes a autovalores maiores que 1;

V) Definição das cargas fatoriais correspondentes à correlação de Pearson entre as variáveis originais e cada um dos fatores;

VI) Observação das comunalidades, de modo a verificar a variância total compartilhada com os fatores extraídos a partir de autovalores maiores que 1;

VII) Elaboração do gráfico para as cargas fatoriais (Favero; Belfiore, 2017).

A PCA é uma técnica estatística amplamente utilizada na modelagem da digestão anaeróbia, especialmente na produção de bio-hidrogênio. Essa abordagem permite a redução da dimensionalidade dos dados, identificando as principais variáveis que contribuem para a variabilidade nos processos de digestão anaeróbia, o que é crucial dada a complexidade e a quantidade de dados gerados em experimentos nessa área.

A aplicação do PCA na digestão anaeróbia possibilita a identificação de padrões ocultos e a eliminação de variáveis redundantes, facilitando a interpretação dos dados e melhorando a eficácia dos modelos preditivos. Estudos têm mostrado que o uso do PCA pode auxiliar na seleção de parâmetros operacionais ideais, como temperatura, pH e carga orgânica, que impactam diretamente a eficiência da produção de bio-hidrogênio. Por exemplo, pesquisas indicam que a análise dos componentes

principais pode revelar como diferentes condições ambientais e de alimentação influenciam as comunidades microbianas envolvidas na digestão, levando a um aumento na produção de biogás e bio-hidrogênio (Dordio *et al.*, 2020; Cai *et al.*, 2018).

Além disso, o PCA também é utilizado em conjunto com técnicas de regressão, tanto linear quanto não linear, permitindo a modelagem mais precisa das relações entre as variáveis operacionais e a produção de bio-hidrogênio. Isso é fundamental para otimizar os processos de digestão anaeróbia e maximizar a produção de energia renovável a partir de resíduos orgânicos (Sulaiman *et al.*, 2021; Mansaray; Kazi, 2020).

Em síntese, a Análise de Componentes Principais é uma ferramenta poderosa na modelagem da digestão anaeróbia de bio-hidrogênio, oferecendo insights valiosos sobre as interações entre variáveis e contribuindo para a otimização de processos, com implicações significativas para a produção de bioenergia sustentável.

3.5.2. Análise de Regressão Linear e Não Linear na Modelagem da Digestão Anaeróbia

A aplicação de análises de regressão linear e não linear na modelagem da digestão anaeróbia tem se mostrado uma estratégia eficaz e fundamental para compreender e otimizar a produção de biogás. Essas ferramentas estatísticas oferecem uma base sólida para prever o comportamento do sistema e identificar fatores críticos que afetam a eficiência do processo, como apontado por estudos recentes (Sulaiman *et al.*, 2021; Mansaray; Kazi, 2020). A aplicação de regressão linear e não linear permite aos pesquisadores e operadores entenderem as complexas interações entre variáveis operacionais, como temperatura, pH e composição do substrato, e seu impacto sobre a eficiência do processo, incluindo a produção de bio-hidrogênio (Kafle; Kim, 2013; Cai *et al.*, 2018).

No caso da regressão linear, é uma técnica frequentemente aplicada para avaliar e prever relações diretas entre variáveis independentes e a produção de bio-hidrogênio, proporcionando modelos simples e interpretáveis que auxiliam na otimização de condições operacionais. Este tipo de análise é especialmente útil para definir como variáveis como temperatura e composição do substrato impactam diretamente a produção de biogás e bio-hidrogênio, permitindo que gestores de biodigestores ajustem variáveis com maior controle sobre o processo (Cai *et al.*, 2018).

Por outro lado, as interações no processo de digestão anaeróbia nem sempre seguem uma relação linear, destacando a importância de métodos não lineares para capturar dinâmicas mais complexas. Técnicas de regressão não linear são mais adequadas para modelar essas relações complexas, oferecendo maior precisão e previsibilidade em cenários onde a interação entre variáveis não é diretamente proporcional (Sulaiman *et al.*, 2021).

Muitos processos biológicos apresentam interações complexas entre variáveis, o que exige modelos de regressão não linear. Estes modelos podem capturar melhor as curvas de crescimento e produção de biogás em condições não estáveis, como demonstrado por Sulaiman *et al.* (2021). A integração de ambas as técnicas permite que operadores ajustem processos e aumentem a eficiência operacional (Kafle; Kim, 2013; Mansaray; Kazi, 2020).

Modelos de regressão não linear são aplicados quando a relação entre variáveis exibe complexidade que uma linha reta não descreve adequadamente. Esse tipo de análise é especialmente relevante para o processo de digestão anaeróbia, onde fatores como a cinética de degradação e as dinâmicas populacionais microbianas criam interações complexas. Sulaiman *et al.* (2021) destacam que os modelos de regressão não linear são essenciais para descrever com precisão a taxa de produção de bio-hidrogênio, incorporando a degradação progressiva dos substratos. Além disso, estudos como o de Kafle e Kim (2013) sugerem que a análise de variáveis ambientais, como umidade, temperatura e carga orgânica, pode proporcionar uma compreensão mais profunda de como esses fatores afetam a eficiência da digestão anaeróbia. Esses modelos permitem explorar e prever a variabilidade do sistema sob diferentes condições operacionais, melhorando o controle do processo e a otimização da produção de biogás.

Na análise de regressão e modelagem de processos de digestão anaeróbia, especialmente na produção de bio-hidrogênio, a validação de modelos depende de pressupostos estatísticos rigorosos, como normalidade dos resíduos, homocedasticidade e ausência de multicolinearidade. Cada um desses aspectos é essencial para garantir a precisão e confiabilidade do modelo.

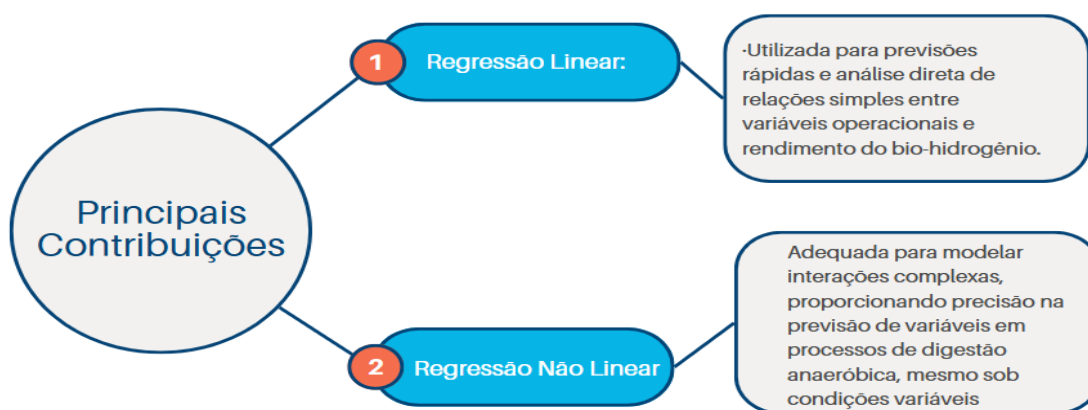
O teste de normalidade é fundamental para determinar a adequação do modelo de regressão, principalmente para a regressão linear, que pressupõe distribuição normal dos erros. Testes como Shapiro-Wilk e Kolmogorov-Smirnov ajudam a avaliar essa suposição, garantindo que a análise estatística seja válida. No

contexto da digestão anaeróbia, a normalidade dos resíduos contribui para a precisão dos modelos preditivos, especialmente quando se utiliza regressão linear para prever produção de biogás e bio-hidrogênio (Cai *et al.*, 2018), e pode assegurar que o modelo pode capturar variações naturais nesse processo de produção (Sulaiman *et al.*, 2021).

Já a homocedasticidade se refere à suposição de que a variância dos resíduos deve ser constante em todo o intervalo das variáveis preditoras. Quando essa condição não é atendida, há heterocedasticidade, o que pode levar a erros de interpretação dos coeficientes. Em estudos sobre digestão anaeróbia, a variabilidade das condições, como a carga orgânica e a temperatura, pode influenciar a consistência do modelo. Identificar e corrigir problemas de heterocedasticidade, por meio de testes como Breusch-Pagan ou White, é crucial para melhorar a precisão da regressão no contexto da produção de bio-hidrogênio (Kafle; Kim, 2013; Cai *et al.*, 2018).

A multicolinearidade, por sua vez, ocorre quando variáveis independentes estão altamente correlacionadas, dificultando a interpretação dos coeficientes e levando a instabilidade no modelo. Em processos de digestão anaeróbia, onde variáveis ambientais frequentemente se correlacionam, a multicolinearidade pode comprometer a análise de quais fatores mais influenciam a produção de bio-hidrogênio. Para mitigar este problema, métodos de avaliação, como o Variance Inflation Factor (VIF), são aplicados para verificar redundâncias e aprimorar a eficácia da modelagem (Mansaray; Kazi, 2020).

Figura 3: Principais Contribuições da Regressão Linear e Não linear no contexto da digestão anaeróbia.



4. METODOLOGIA

Visando observar possíveis correlações entre distintas variáveis do processo anaeróbio e verificar a aplicação de modelos lineares e não-lineares na predição da produção de bio-hidrogênio, foram utilizados dados coletados em pesquisas prévias realizadas no Laboratório de Controle Ambiental (LCA), do Centro de Tecnologia da UFAL.

Os trabalhos foram usados como referência por apresentarem padronização quanto aos métodos laboratoriais na análise dos parâmetros, tais como:

- pH, DQO e concentração de sólidos: Standard methods for the examination of water and wastewater (APHA, 2017);
- Carboidratos: Método colorimétrico para determinação de açúcares e substâncias relacionadas (Dubois *et al.*, 1956);
- Concentração de ácidos voláteis, álcool e composição do biogás: Cromatografia usando o equipamento GC-2010 Shimadzu (Adorno *et al.*, 2014).

Desta forma, foram selecionadas as pesquisas voltadas à produção biológica de hidrogênio em modo contínuo (utilizando reatores de leito fluidificado) ou em batelada, a partir do uso de diferentes resíduos gerados em Alagoas.

4.1. Determinação das variáveis e Agrupamento de dados

Para a organização dos dados e realização da meta-análise, foram extraídas informações referentes ao rendimento de hidrogênio e produção acumulada do mesmo após o processo de digestão anaeróbia. Esta fase envolve a análise exploratória multivariada para diminuir o volume de informações das variáveis originais para um grupo mais restrito de variáveis estatísticas (Jolliffe; Cadima, 2016). As interações foram avaliadas entre o grupo de estudos através de seus componentes principais, assim foi possível inferir semelhanças ou discrepâncias com base na distância entre os pontos amostrais.

As variáveis composicionais, operacionais e respostas foram separadas de acordo com o tipo de operação do reator (batelada ou contínuo), como apresentado na Tabela 5:

Tabela 5: Variáveis composicionais, operacionais e respostas.

Reator	Variáveis composicionais	Variáveis operacionais	Variáveis respostas
Batelada	Concentração de Sólidos Voláteis; Volume de Inóculo; Razão substrato e inóculo (SI) Remoção de DQO; Remoção de carboidrato;	pH inicial; Agitação; Tempo de digestão anaeróbia (dias);	Produção acumulada de Hidrogênio;
Reator	Variáveis composicionais	Variáveis operacionais	Variáveis respostas
Contínuo	Demanda Química de Oxigênio inicial (DQO inicial); Carboidrato inicial (CHT inicial);	pH inicial; Temperatura; Taxa de carregamento orgânico (TCO); Tempo de Detenção Hidráulica (TDH); Tempo de digestão anaeróbia (dias);	Rendimento de Hidrogênio;

Fonte: Autora (2024)

4.2. Aspectos gerais dos estudos

Nesta revisão sistemática, um total de 25 registros foram encontrados nas bases de dados Scopus e Sucupira. Após a triagem e avaliação manual, 13 arquivos foram descartados por não estarem em conformidade com os critérios pré-estabelecidos. No final, 12 documentos foram avaliados minuciosamente e analisados estatisticamente. Esses documentos apresentaram experimentos usando resíduos agroindustriais da região semiárida do Estado de Alagoas e inóculo na digestão anaeróbia. A tabela 6 apresenta características gerais dos estudos selecionados.

Tabela 6: Características gerais dos estudos.

Objetivo	Método (Reator)	Nº Experimentos	Inóculo	Referência
Avaliar as variáveis de processo envolvidas na produção de bio-hidrogênio em reatores anaeróbios em batelada a partir de frutas e vegetais residuais oriundos de um centro de abastecimento.	Batelada	16	Lodo Anaeróbio	Lima, P.G. (2020)
Avaliar a produção de hidrogênio em reator batelada por fermentação	Batelada	8	Lodo Anaeróbio	Balbino, E.S. (2020)

anaeróbia utilizando como substrato o hidrolisado de palha da soja.				
Avaliar a eficiência de um reator anaeróbio em relação à produção de hidrogênio a partir de casca de banana como substrato.	Batelada	16	Lodo Anaeróbio	Soares, J.D.M. (2019)
Avaliar a produção de hidrogênio a partir da vinhaça da cana-de-açúcar como fonte de carbono em diferentes concentrações.	Batelada	4	Lodo Anaeróbio	Telles, N.L.S. (2015)
Avaliar a produção de hidrogênio a partir do resíduo de cerveja coletado do tanque de fermentação como fonte de carbono, procurando avaliar os efeitos da Demanda Química de Oxigênio – DQO, tipo de inóculo e agitação (sem agitação e 120 rpm)	Batelada	8	Fermentação natural e lodo anaeróbio	Telles, N.L.S. (2023)
Avaliar a produção de biocombustíveis e produtos de valor agregado em um reator anaeróbio de leito fluidificado através da cofermentação de melaço de cana-de-açúcar e glicerol bruto.	Contínuo (RALF)	4	Lodo Anaeróbio	Pereyra, D.L.A.D. (2019)
Avaliar o efeito do tempo de detenção hidráulica em um reator anaeróbio de leito fluidificado (RALF) com soro de leite como fonte de carbono em temperatura Mesofílica.	Contínuo (RALF)	5	Lodo Anaeróbio	Marques, T.D. (2018)
Avaliar a produção simultânea de hidrogênio e metano em um sistema combinado acidogênico-metanogênico, utilizando reatores anaeróbios de leito fluidizado rearranjados em série e alimentados com vinhaça da cana-de-açúcar como fonte de carbono.	Contínuo (RALF)	3	Lodo Anaeróbio	Gois, G.N.S.B. (2017)
Avaliar a produção de hidrogênio a partir de um reator anaeróbio de leito fluidificado operado continuamente e alimentado com o substrato melaço da cana-de-açúcar.	Contínuo (RALF)	4	Lodo Anaeróbio	Chaves, T.C. (2017)
Avaliar a produção de hidrogênio em um reator anaeróbio de leito fluidificado utilizando efluente de laticínio como substrato.	Contínuo (RALF)	4	Lodo Ativado	Silva, A.N. (2015)
avaliar a eficiência da produção de hidrogênio, em reator anaeróbio de leito fluidificado, a partir de resíduos da suinocultura, suplementado com sacarose.	Contínuo (RALF)	9	Lodo Ativado	Santos, E.C.L. (2014)
Avaliar a produção biológica de hidrogênio a partir da associação entre a manipueira e esgoto sanitário em dois reatores de leito fluidificado (RALF 1 e 2).	Contínuo (RALF)	11	Lodo Ativado	Vilela, L.M.B. (2013)

Avaliar a produção de hidrogênio em reator anaeróbio de leito fluidificado, alimentado com água residuária do processamento da mandioca – a manipueira – como fonte de carbono, analisando a influência de temperaturas termofílicas.	Contínuo (RALF)	3	Lodo Ativado	Tavares, T.G.A. (2016)
---	-----------------	---	--------------	------------------------

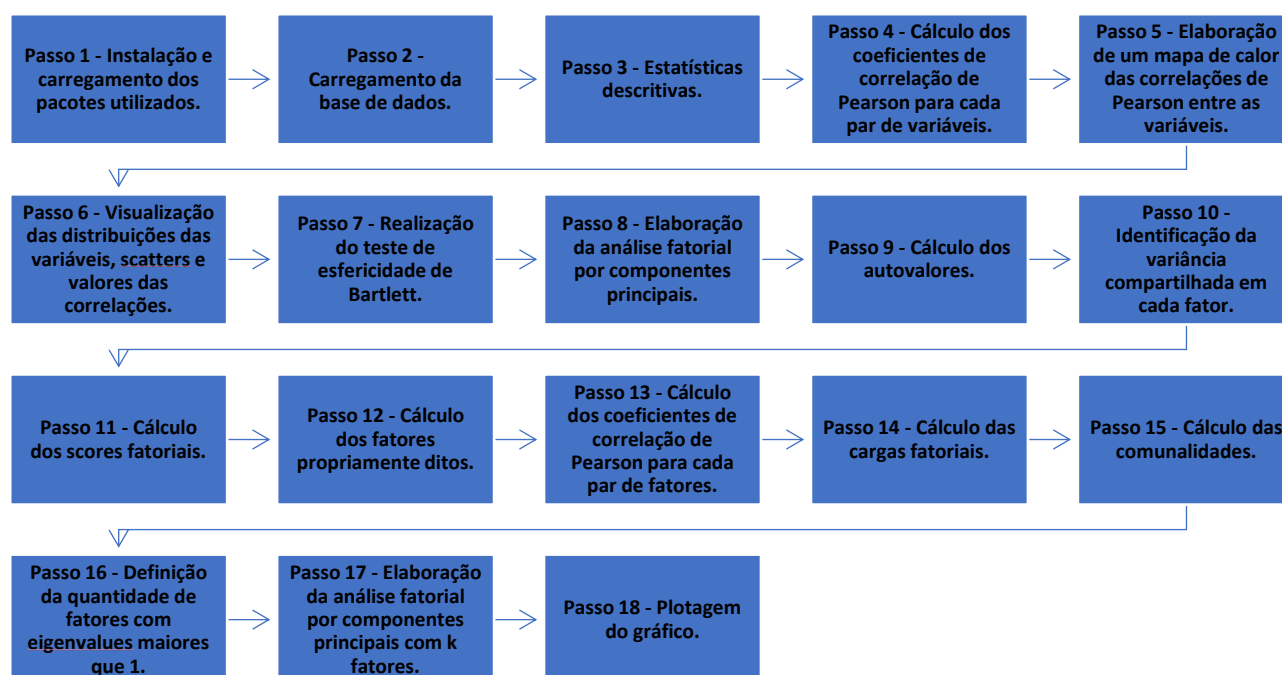
Fonte: Autora (2024)

4.3.1. Análise de Componentes Principais (PCA)

Os dados selecionados foram processados pelo software estatístico livre R Project, para análise estatística multivariada. A análise de componentes principais foi realizada a partir da matriz de correlação com os dados auto escalonados.

As etapas de implementação podem ser resumidas de acordo com a Figura 4:

Figura 4: Etapas Da Implementação Da PCA No Software R.



Fonte: Autora (2024)

A interpretação dos componentes principais analisou a influência de cada variável característica sobre cada componente, nesse caso sobre cada substrato, de acordo com o grau de importância, com o objetivo de fazer comparações entre os indivíduos e/ ou agrupá-los, e assim calcular para cada indivíduo os scores como para cada componente. De acordo com Alburquerque *et al.*, (2016), os critérios de seleção dos componentes principais (PC) consistem em incluir somente os componentes com

autovalores maiores que 1 ou componentes com variância acumulativa acima de 70%.

A combinação da digestão anaeróbia com ferramentas estatísticas e de modelagem multivariada, como a PCA e modelos de regressão, tem se mostrado uma abordagem promissora para maximizar a eficiência da produção de biogás. PCA é comumente usada para simplificar a análise de dados de digestores anaeróbicos, reduzindo a dimensionalidade dos dados ao identificar os principais componentes que influenciam o processo, como temperatura, pH, inóculo, e tempo de retenção, etc. Essa técnica permite entender melhor a influência dessas variáveis inter-relacionadas e ajuda a eliminar redundâncias nos dados (Kafle; Kim, 2013; Sulaiman *et al.*, 2021).

4.3.2. Análise de Regressão Múltipla Linear e Não Linear

Um modelo de regressão linear assume que uma função de regressão é linear nas variáveis de entrada. Esses modelos são simples e muitas vezes fornecem uma descrição adequada e interpretável de como variáveis explicativas afetam os valores previstos. Em geral, o objetivo da regressão linear é encontrar os parâmetros α (coeficiente linear) e β (coeficiente angular) para os quais o termo de erro é minimizado. A técnica mais usual de estimação desse modelo é o método dos mínimos quadrados, que foi utilizado nesse estudo (Hastie *et al.*, 2009).

No entanto, há situações em que as relações entre variáveis podem se manifestar de várias formas funcionais não lineares, as quais devem ser consideradas durante a estimação de modelos de regressão para uma compreensão mais precisa do comportamento dos diversos fenômenos. Nestes casos, a partir do modelo de regressão linear, é possível derivar um modelo de regressão não linear através da transformação de Box-Cox. Esse método implica na substituição da variável resposta Y por parâmetros de transformação, conforme determinada forma funcional (linear, semilogarítmica, logarítmica, inversa, quadrática ou cúbica) (Fávero; Belfiore, 2017). Portanto, no caso da Regressão Não-Linear, foi utilizado o método de transformação de Box-Cox.

Segundo Fávero e Belfiore (2017) “a significância estatística geral do modelo e de seus parâmetros estimados não é determinada pelo R^2 , mas sim através de testes estatísticos apropriados”. Desta forma, a adequação dos modelos de regressão foi avaliada a partir dos pressupostos apresentados na tabela 7.

Tabela 7: Testes Estatísticos.

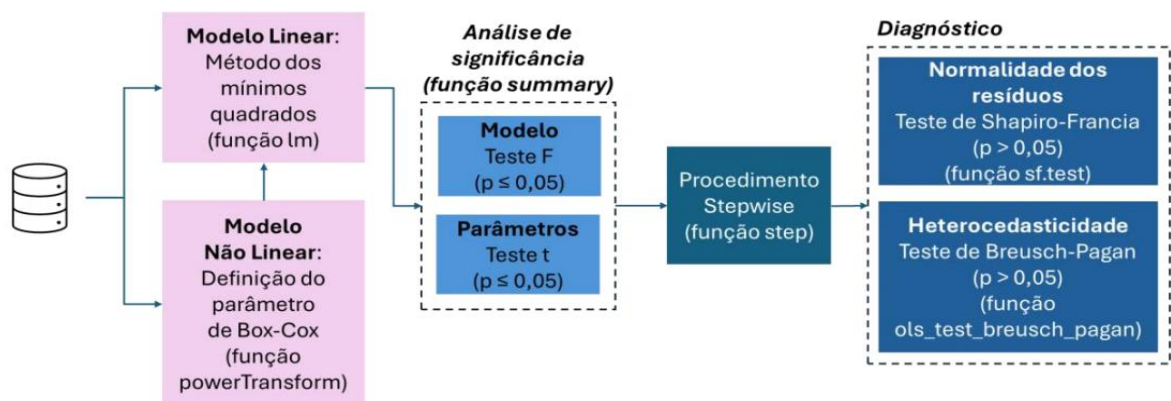
Pressuposto	Verificação	Diagnóstico	Interpretação
Significância geral do modelo	Teste F	valor p de	Para o modelo ser considerado significativo, pelo menos um coeficiente angular (β), referente a uma variável explicativa, deve ser estatisticamente diferente de zero.
Significância de cada parâmetro	Testes t de Student	p-valor deve ser $\leq 0,05$	Para os parâmetros α e β do modelo serem considerados significativos, eles devem ser estatisticamente diferentes de zero.
Resíduos normalmente distribuídos	Teste de Shapiro-Francia	p-valor deve ser $> 0,05$	Quando os resíduos não são normalmente distribuídos, pode-se interpretar que houve a especificação incorreta do modelo e variáveis explicativas relevantes não foram incluídas na modelagem.
Ausência de multicolinearidade	Matriz de correlação	Se a Tolerance for muito baixa e VIF muito alta, há problemas de multicolinearidade	É identificada, indica que existem correlações muito elevadas entre variáveis explicativas ou o número de observações do banco de dados foi insuficiente.
Ausência de heterocedasticidade	Teste de Breusch-Pagan	p-valor deve ser $> 0,05$	A ocorrência de heterocedasticidade significa que a variância dos resíduos não foi constante, gerada provavelmente por erros de especificação quanto à forma funcional ou quanto à omissão de variável.

Fonte: Adaptado de Fávero e Belfiore (2017)

4.3.2.1. Modelagem no R

A Figura 5 resume as etapas da modelagem supervisionada no R, utilizando o método dos mínimos quadrados para análise de regressão linear e não linear, utilizando a transformação de Box-Cox.

Figura 5: Etapas da modelagem supervisionada no R.



Fonte: Peiter (2024)

5. RESULTADOS E DISCUSSÕES

5.1. Análise Fatorial por Componentes Principais (Reatores operados em batelada)

5.1.1. Estatísticas descritivas

Como descrevem Bro e Smilde (2014), as estatísticas descritivas ajudam a verificar se os dados apresentam variância suficiente para justificar a aplicação da PCA e se há variáveis com distribuições anômalas que possam distorcer os componentes. Jolliffe e Cadima (2016) reforçam que, ao avaliar essas estatísticas, é possível identificar a necessidade de normalização ou padronização, especialmente quando variáveis apresentam escalas de medição diferentes, o que pode influenciar o resultado da PCA. Além disso, variáveis com valores extremos podem impactar os componentes principais, e a análise das estatísticas descritivas auxilia na decisão de transformações adicionais para adequar o modelo.

A inspeção das estatísticas descritivas, portanto, fornece insights sobre o preparo necessário dos dados para que a PCA seja aplicada de maneira eficaz e os componentes principais representem a estrutura subjacente dos dados.

As estatísticas descritivas apresentadas para cada variável da Tabela 8 fornecem uma visão geral da distribuição dos dados antes da PCA. Para uma análise robusta das estatísticas descritivas desta amostra de dados sobre estudos em reatores em batelada, observamos medidas fundamentais como média, mediana, valores mínimos e máximos, com base em variáveis operacionais que impactam diretamente a produção acumulada de hidrogênio (HY). Essas estatísticas fornecem uma visão inicial sobre a distribuição dos dados e seus pontos centrais e extremos, o que é essencial para análise subsequente em modelos de regressão e PCA. (Favero; Belfiore, 2017).

Tabela 8: Estatísticas descritivas (Modo Batelada).

Variável	Min	1st Qu,	Median	Mean	3rd Qu,	Max
pH_inicial	-4,00	5,48	5,57	5,93	7,00	7,70
SVI	0,51	0,51	2,90	6,21	7,25	21,75
VI	0,05	0,10	0,10	0,24	0,25	0,75
Dias	7,00	8,00	20,00	22,08	33,00	69,00
Agitação	0,00	-120,00	-120,00	-101,50	-120,00	-120,00
SI	0,33	3,00	9,00	7,36	9,00	19,00
HY	0,38	64,80	98,38	126,46	153,34	433,40

Fonte: Adaptado do software R (2024)

Com valores entre 4,0 e 7,7, observa-se que o pH inicial da amostra tem um valor médio de 5,93, indicando uma leve tendência à faixa ácida, que é muitas vezes propícia para processos de fermentação anaeróbia (Zhen *et al.*, 2015). Com amplitude entre 0,51 e 21,75, o SVI possui uma média de 6.21. A mediana mais baixa que a média indica assimetria e possivelmente valores extremos, que podem representar condições experimentais específicas.

VI (Volume de Inóculo) possui valores entre 0,05 e 0,75, o volume médio de inóculo é de 0,24. Este volume é crucial para determinar a carga orgânica e seu impacto nos processos de produção de hidrogênio. A variável “Dias” apresenta média de 22,08 dias, variando de 7 a 69 dias, o que mostra a duração dos processos.

A variável “Agitação”, com média de 101,5 rpm, reflete os diferentes regimes de mistura nos reatores, variando de 0 a 120 rpm. Com valores entre 0,33 e 19,0 e média de 7,35, a razão entre substrato e inóculo reflete a proporção de material orgânico disponível. A produção acumulada de hidrogênio possui média elevada (126,46), embora exista uma alta variabilidade, com o menor valor registrado em 0,38 e o maior em 433,40.

Essas estatísticas descritivas mostram uma variabilidade significativa em algumas variáveis, como HY e SVI, recomendando uma verificação de normalidade e homogeneidade para garantir a robustez dos modelos de regressão subsequentes (Favero; Belfiore, 2017).

5.1.2. Coeficientes de correlação de Spearman para cada par de variáveis

Diferente do coeficiente de Pearson, que assume linearidade e distribuição normal, o coeficiente de Spearman considera apenas as ordens dos valores, sendo apropriado para variáveis ordinais ou para situações em que a relação entre variáveis não é linear. Essa medida varia de -1 a 1, onde valores próximos a -1 indicam uma forte correlação negativa, valores próximos a 1 indicam uma forte correlação positiva e valores próximos a 0 sugerem pouca ou nenhuma correlação (Corder; Foreman, 2014).

5.1.2.1. Matriz de correlações

A matriz de correlações em uma PCA mostra as correlações entre cada par de variáveis, sendo útil para entender o agrupamento e a contribuição de cada variável para a construção dos componentes principais (Jolliffe; Cadima, 2016). A análise da

matriz de correlações entre variáveis operacionais nos reatores em batelada usando o coeficiente de correlação de Spearman permite avaliar as interdependências não lineares entre elas, uma prática comum e bem recomendada em estudos envolvendo digestão anaeróbia devido à natureza não-paramétrica do método (Favero; Belfiore, 2017).

Tabela 9: Matriz de Correlação de Spearman (Modo Batelada).

	ph_inicial	SVI	VI	Dias	Agitação	SI	HY
ph_inicial	1,000	0,069	0,129	-0,569	0,341	-0,129	-0,153
SVI	0,069	1,000	0,942	0,388	0,117	-0,942	0,417
VI	0,129	0,942	1,000	0,309	0,239	-1,000	0,446
Dias	-0,569	0,388	0,309	1,000	-0,382	-0,309	0,241
Agitação	0,341	0,117	0,239	-0,382	1,000	-0,239	0,348
SI	-0,129	-0,942	-1,000	-0,309	-0,239	1,000	-0,446
HY	-0,153	0,417	0,446	0,241	0,348	-0,446	1,000

Fonte: Adaptado do Software R (2024).

Ao analisar a matriz de correlação de Spearman, observa-se a interação entre variáveis operacionais e a produção acumulada de hidrogênio (HY), o que é essencial para identificar padrões antes de aplicarmos a PCA. Abaixo, as correlações mais relevantes:

- pH_inicial e Dias (-0,569): A correlação negativa moderada sugere que, à medida que o pH inicial aumenta, o tempo de retenção hidráulica (Dias) tende a diminuir.
- SVI e VI (0,942): A alta correlação positiva entre SVI (concentração de sólidos voláteis do inóculo) e VI (volume de inóculo) sugere que esses parâmetros são ajustados de maneira proporcional, possivelmente para manter uma quantidade constante de biomassa ativa no reator.
- VI e SI (-1,000): A correlação negativa perfeita indica uma dependência linear e inversa entre VI e SI.
- SI e HY (-0,446): A correlação negativa moderada entre a razão substrato/inóculo (SI) e a produção de hidrogênio (HY) sugere que uma razão SI mais alta pode estar associada a uma produção menor de hidrogênio.

5.1.2.2. Matriz com p-valor dos coeficientes

Segundo Favero e Belfiore (2017) os p-valores são cruciais para determinar a significância estatística das correlações observadas na matriz de correlação de Spearman, um p-valor abaixo de 0,05 é geralmente considerado indicativo de uma correlação estatisticamente significativa, e o valor de NA é utilizado quando não há correlação possível entre a variável e ela mesma.

Tabela 10: Matriz com p-valor dos coeficientes (Modo Batelada).

	ph_inicial	SVI	VI	Dias	Agitação	SI	HY
pH_inicial	NA	0,626	0,363	0,000	0,013	0,363	0,278
SVI	0,626	NA	0,000	0,004	0,407	0,000	0,002
VI	0,363	0	NA	0,026	0,088	0,000	0,001
Dias	0,000	0,004	0,026	NA	0,005	0,026	0,086
Agitação	0,013	0,407	0,088	0,005	NA	0,088	0,011
SI	0,363	0,000	0,000	0,026	0,088	NA	0,001
HY	0,278	0,002	0,001	0,086	0,011	0,001	NA

Fonte: Adaptado do software R (2014).

Analisando a matriz de p-valores para as correlações entre as variáveis operacionais e composicionais, observa-se evidências significativas para certas relações, que auxiliam na interpretação de interações potencialmente importantes no contexto da produção de hidrogênio em reatores de batelada:

- pH Inicial e Dias ($p = 0,00001$);
- pH Inicial e Agitação ($p = 0,01333$);
- SVI e VI ($p < 0,00001$);
- SVI e SI ($p < 0,00001$);
- VI e HY ($p = 0,00091$);
- Dias e Agitação ($p = 0,00519$);
- SI e HY ($p = 0,00091$);

5.1.2.3. Mapa de Calor das Correlações

De acordo com Fávero e Belfiore (2017), os mapas de calor das correlações de Spearman desempenham um papel crucial na Análise de Componentes Principais (PCA) ao permitir a visualização clara das relações entre variáveis antes de aplicar a PCA. Os mapas de calor ajudam a identificar padrões e estruturas nos dados,

destacando variáveis que estão fortemente correlacionadas, o que pode influenciar as escolhas de quais variáveis incluir na análise.

Além disso, Fávero e Belfiore (2017) ressaltam que a aplicação de métodos não paramétricos, como o teste de Spearman, é vantajosa em muitos cenários, pois não requer que os dados atendam a suposições de normalidade, tornando as análises mais robustas em contextos de dados reais, onde essas suposições muitas vezes não são cumpridas.

Figura 6: Mapa de calor das correlações de Spearman entre as variáveis.



Fonte: Software R (2024).

A análise das correlações entre variáveis operacionais e de produção de hidrogênio em reatores de batelada é crucial para entender as interdependências que afetam o desempenho do processo. Usando a correlação de Spearman, observa-se que algumas variáveis possuem fortes relações entre si, o que pode impactar diretamente a eficiência e o controle do processo de digestão anaeróbia para produção de hidrogênio. Abaixo algumas observações detalhadas, baseadas no mapa de calor.

- A concentração de sólidos voláteis do inóculo (SVI) apresenta uma correlação positiva muito forte com o volume de inóculo (VI), com um coeficiente de 0,9423. Isso indica que, à medida que o volume de inóculo aumenta, a concentração de sólidos voláteis também tende a aumentar. Esse comportamento já foi relatado em estudos anteriores, que evidenciam que a quantidade de inóculo influencia diretamente a concentração de sólidos

voláteis, afetando a eficiência do processo de digestão anaeróbia (FAVERO; BELFIORE, 2021).

- SI (Razão Substrato/Inóculo) e VI (Volume de Inóculo): Há uma correlação negativa muito forte entre SI e VI (-1,0000), indicando que o aumento do volume de inóculo reduz a razão substrato/inóculo. Essa relação inversa sugere que a modulação do volume de inóculo é uma estratégia eficaz para controlar a razão SI, fator crítico na otimização da produção de hidrogênio (SILVA *et al.*, 2019).
- A correlação negativa entre o número de dias de operação e o pH inicial (-0,5695) indica que, com o aumento dos dias de operação, há uma tendência para a redução do pH inicial. Esse comportamento pode estar relacionado à acumulação de ácidos orgânicos ao longo do tempo, resultando na diminuição do pH, o que pode influenciar a atividade microbiana e a eficiência da produção de hidrogênio (HALLAM, 2016).

A matriz de p-valores complementa a matriz de correlações ao indicar a significância estatística das correlações observadas, o mapa de calor ajuda na visualização dessas correlações. Por exemplo, as correlações entre VI e SVI, bem como entre SI e VI, são significativas, reforçando a interpretação de que essas interações são relevantes no contexto operacional do processo de digestão anaeróbia. Esse resultado suporta a recomendação de estudos prévios, como o de Favero e Belfiore (2021), que destacam a necessidade de otimização desses parâmetros para melhorar o rendimento de hidrogênio.

5.1.3. *Elaboração a Análise Fatorial Por Componentes Principais*

5.1.3.1. *Teste de esfericidade de Bartlett*

O teste de esfericidade de Bartlett é usado para verificar a adequação de dados para análise fatorial, especialmente em Análise Fatorial por Componentes Principais (PCA). Os autovalores e a proporção da variância explicada para cada componente principal são essenciais para identificar quantos componentes devemos reter na análise de componentes principais (PCA). A interpretação dos autovalores e das variâncias é frequentemente orientada pela regra de Kaiser, que sugere a retenção dos componentes com autovalores maiores que 1 (Favero *et al.*, 2009).

5.1.3.2. Eigenvalues (autovalores) e Variância compartilhada pelas variáveis originais para a formação de cada fator

Na análise de componentes principais (PCA) apresentada, são calculados autovalores e variâncias compartilhadas para entender a contribuição de cada componente principal na variabilidade total dos dados.

Tabela 11: Autovalores e Variância compartilhada pelas variáveis (Modo Batelada).

	PC1	PC2	PC3	PC4	PC5	PC6
Autovalores	2.815	1.710	0.903	0.304	0.263	0.006
Prop. da Variância	0.469	0.285	0.151	0.051	0.044	0.001
Prop. da Variância Acumulada	0.469	0.754	0.905	0.955	0.999	1.000

Fonte: Software R (2024).

➤ Autovalores

- PC1 possui um autovalor de 2,815 e explica 46,9% da variância total. Este é o componente que mais contribui para explicar a variabilidade dos dados, indicando que pode representar uma combinação das variáveis com maior poder explicativo sobre o fenômeno estudado.
- PC2 tem um autovalor de 1,710 e contribui com 28,5% da variância. Junto com o PC1, os dois primeiros componentes explicam 75,4% da variância acumulada. o que é considerado um valor satisfatório para reduzir a dimensionalidade dos dados sem perder muita informação (Jolliffe, 2002).

➤ Variância Acumulada

A variância acumulada é de 90,5% até o terceiro componente e atinge 100% com todos os componentes. Normalmente, recomenda-se reter os componentes que explicam pelo menos 70-80% da variância total para garantir uma representação robusta dos dados (Jolliffe, 2002). Assim, os componentes PC1 e PC2, e possivelmente o PC3, parecem ser suficientes para uma análise fatorial eficaz, pois juntos explicam mais de 90% da variância.

Essa análise indica que a maior parte das informações dos dados originais pode ser capturada com apenas dois ou três componentes principais, o que é vantajoso para simplificação e interpretação.

5.1.3.3. Scores Fatoriais de cada Variável em cada Componente

Os scores fatoriais refletem o peso de cada variável em cada componente e permitem a classificação de amostras com base nas características capturadas pelos componentes principais. O uso dos scores fatoriais é essencial em estudos para prever ou classificar amostras, permitindo uma análise mais simplificada dos dados originais (Jolliffe, 2002).

Tabela 12: Score Fatoriais (Modo Batelada).

	PC1	PC2
pH_inicial	-0.009	0.334
SVI	0.345	0.041
VI	0.343	0.051
Dias	0.144	-0.481
Agitação	0.069	0.486
SI	-0.305	-0.024

Fonte: Software R (2024).

Para analisar os scores fatoriais, observa-se o peso de cada variável nos primeiros dois componentes principais (PC1 e PC2). A interpretação desses scores pode indicar quais variáveis possuem mais influência em cada componente, permitindo identificar padrões nos dados.

PC1 pode estar mais relacionado a fatores composicionais do inóculo, como o SVI e o VI, indicando que a composição inicial do inóculo tem um papel relevante na produção. PC2 reflete as condições operacionais, como agitação e tempo de incubação, que também afetam a eficiência do processo de digestão anaeróbia.

5.1.3.4. Cargas Fatoriais de cada variável com os componentes principais

As cargas fatoriais indicam o grau de correlação de cada variável com os componentes principais. Cargas fatoriais acima de 0,5 geralmente são consideradas significativas (Hair *et al.*, 2014).

Tabela 13: Cargas Fatoriais (Modo Batelada).

	PC1	PC2
pH_inicial	-0.024	0.572
SVI	0.971	0.070
VI	0.964	0.087
Dias	0.406	-0.823
Agitação	0.194	0.832
SI	-0.860	-0.040

Fonte: Software R (2024).

As cargas fatoriais fornecem informações sobre como cada variável está associada aos componentes principais, permitindo uma interpretação sobre quais fatores estão influenciando mais significativamente cada componente. A análise das cargas fatoriais para PC1 e PC2 nos dados dos reatores em batelada revela alguns padrões importantes.

PC1 reflete um fator composicional, fortemente influenciado pela concentração e volume do inóculo, sendo essencial na configuração inicial dos reatores para otimizar a produção de hidrogênio. Isso se alinha com estudos sobre digestão anaeróbia, que destacam a importância do controle dos parâmetros iniciais do inóculo (Favero; Belfiore, 2021). PC2 representa fatores operacionais, como agitação e tempo, que são variáveis ajustáveis durante o processo e impactam diretamente na estabilidade e produção. Estudos como o de Belfiore e Silva (2019) demonstram que a otimização desses parâmetros pode melhorar a eficiência da produção de biocombustíveis.

5.1.3.5. Comunalidades

As comunalidades representam a variância compartilhada por cada variável com os componentes selecionados e indicam o quão bem cada variável é explicada pelos fatores retidos. Comunalidades próximas de 1 indicam que a maior parte da variância de uma variável é explicada pelos componentes selecionados (Tabachnick; Fidell, 2014).

Tabela 14: Comunalidades (Modo Batelada).

comunalidades	
pH_inicial	0.33
SVI	0.95
VI	0.94
Dias	0.84
Agitação	0.73
SI	0.74

Fonte: Software R (2024).

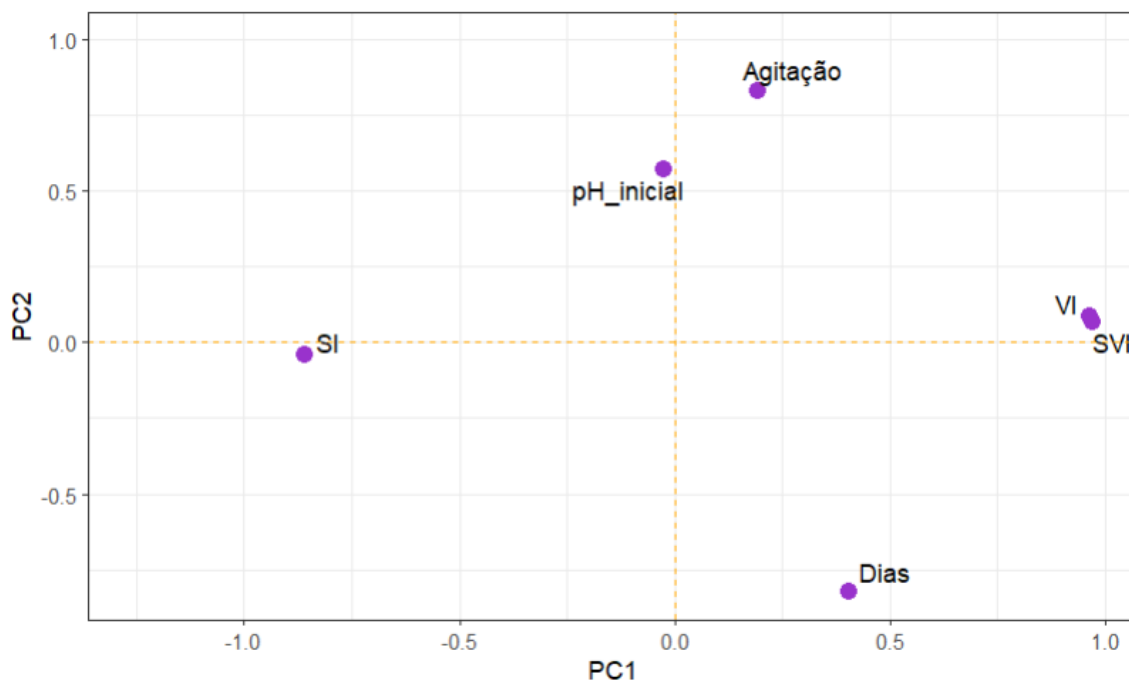
Segundo Favero e Belfiore (2021), a presença de altos valores de comunalidade em variáveis relacionadas ao inóculo reforça sua importância na configuração dos sistemas de digestão anaeróbia.

A análise das comunalidades indica que as variáveis SVI e VI desempenham papéis cruciais no modelo, capturando a maior parte da variância. A variável Dias também contribui significativamente, enquanto Agitação e SI apresentam um efeito moderado. O pH inicial mostra-se menos representativo nos componentes principais, sugerindo que seu impacto é mais limitado.

5.1.3.6. Gráfico das Cargas Fatoriais

O gráfico das cargas fatoriais nos dois primeiros componentes principais (PC1 e PC2) é uma representação visual da relação entre as variáveis e os fatores extraídos. Esse tipo de gráfico permite uma interpretação mais clara sobre quais variáveis estão mais fortemente associadas aos componentes principais, conforme destacado em análises de PCA em estudos de digestão anaeróbia (Favero; Belfiore, 2021).

Figura 7: Gráfico dos dois primeiros Componentes Principais (Modo Batelada).



Fonte: Software R (2024).

O gráfico das cargas fatoriais mostra que o PC1 captura variáveis relacionadas à composição do inóculo, como SVI e VI, com influência oposta da razão substrato-inóculo (SI). Já o PC2 reflete o efeito de variáveis operacionais, como pH inicial, Agitação e Dias. Essa separação reforça a ideia de que, para otimizar a produção de hidrogênio, é crucial considerar tanto os fatores composicionais quanto os operacionais, com enfoque nas variáveis de inóculo no contexto dos estudos de digestão anaeróbia (Soares *et al.*, 2022; Belfiore; Silva, 2019).

5.2. Análise de Regressão para modo Batelada

5.2.1. Estatísticas Univariadas

A análise univariada é um passo inicial essencial para entender a distribuição e a variabilidade dos dados, e ajuda a identificar possíveis valores extremos que podem influenciar nas etapas subsequentes da análise multivariada (Favero; Belfiore, 2019).

Tabela 15: Estatísticas Univariadas (Modo Batelada).

Variável	Min	1st Qu.	Median	Mean	3rd Qu.	Max
pH_inicial	-4,000	5,478	5.565,000	5,926	7,000	7,700
SVI	0,507	0,507	2,900	6,206	7,250	21,750
VI	0,050	0,100	0,100	0,238	0,250	0,750
Dias	7,000	8,000	20,000	22,080	3,300	69,000
Agitação	0,000	-120,000	-120,000	-101,500	-120,000	-120,000
SI	0,330	3,000	9,000	7,358	9,000	19,000
HY	0,379	64,800	98,380	126,461	153,335	433,400

Fonte: Adaptado do software R (2024)

Análise:

- pH_inicial: Observa-se que o pH inicial varia de 4 a 7,7, com média de 5,93, indicando uma leve acidez média nos dados. O pH é um dos principais fatores que influenciam a digestão anaeróbia, impactando diretamente a atividade das comunidades microbianas envolvidas no processo fermentativo (Silva *et al.*, 2018). Valores próximos à neutralidade tendem a favorecer a produção de hidrogênio, enquanto pH muito ácido pode inibir a atividade de bactérias hidrogenogênicas (Amaral; Steinmetz; Kunz, 2017).
- SVI (Concentração de Sólidos Voláteis do Inóculo): A média de 6,21 indica a presença significativa de material orgânico no inóculo, essencial para a estabilidade da digestão anaeróbia. A concentração de sólidos voláteis influencia diretamente a taxa de degradação da matéria orgânica e a eficiência na produção de biogás (Souza, 2014).
- VI (Volume de Inóculo): A média de 0,238 mostra uma tendência de uso de pequenas quantidades de inóculo, variando entre 0,05 e 0,75. A proporção entre inóculo e substrato pode afetar significativamente a cinética do processo, uma vez que baixos volumes podem reduzir a diversidade microbiana e, consequentemente, comprometer a eficiência da digestão anaeróbia (Souza, 2016).
- Dias: A variação no número de dias (7 a 69) reflete diferentes tempos de retenção hidráulica nos experimentos. Tempos de retenção mais longos permitem uma conversão mais eficiente da matéria orgânica, enquanto tempos mais curtos podem comprometer a estabilidade do sistema (Souza, 2014; Amaral; Steinmetz; Kunz, 2017).
- A agitação e a razão substrato/inóculo (SI) também mostram variação

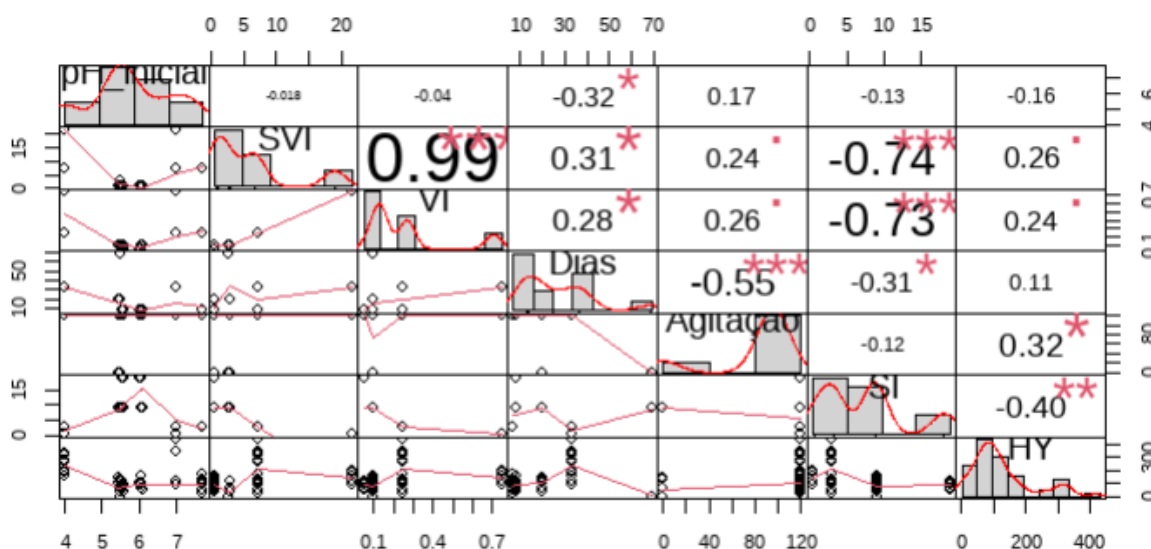
substancial entre os experimentos. A agitação adequada melhora a transferência de massa e a homogeneização do substrato, fatores que impactam diretamente na produção de hidrogênio (Silva *et al.*, 2018).

- SI (Razão Substrato/Inóculo): A média de 7,36 e o valor máximo de 19 mostram a proporção do substrato em relação ao inóculo, influenciando a produção de hidrogênio. Essa relação é fundamental na digestão anaeróbia, pois influencia a disponibilidade de nutrientes e a competição entre microrganismos fermentativos e metanogênicos (Amaral; Steinmetz; Kunz, 2017). Valores muito elevados podem causar acidificação excessiva, reduzindo a produção de hidrogênio (Silva *et al.*, 2018).
- HY (Produção Acumulada de Hidrogênio): Com uma média de 126,46, os dados indicam que o rendimento acumulado de hidrogênio é variável, com um máximo de 433,4.

As estatísticas univariadas fornecem uma visão inicial das características dos dados, permitindo observar possíveis padrões e variabilidades antes de aplicar modelos mais complexos. Em estudos de digestão anaeróbia, a compreensão desses valores é essencial para otimizar as condições operacionais e maximizar a produção de hidrogênio.

5.2.1.1. Estudo das Correlações

Figura 8: Gráfico de correlações (Modo Batelada).



Fonte: Autora (2024)

A análise das correlações entre as variáveis envolvidas nos estudos de produção de hidrogênio em reatores batelada revela relações importantes entre os parâmetros operacionais e composicionais.

- Correlação entre SVI e VI: Há uma correlação extremamente alta e significativa ($r = 0,99$) entre a concentração de sólidos voláteis do inóculo (SVI) e o volume de inóculo (VI).
- Correlação negativa entre SVI e SI: A relação inversa entre SVI e a razão substrato/inóculo (SI) ($r = -0,74$) sugere que aumentos na concentração de sólidos voláteis estão associados a uma redução na razão substrato/inóculo.
- Correlação negativa entre Dias e Agitação: A correlação entre o número de dias e a agitação é negativa e significativa ($r = -0,55$), sugerindo que períodos de operação mais longos tendem a ocorrer com menor agitação.
- Correlação entre SI e HY: A razão substrato/inóculo (SI) apresenta uma correlação negativa com a produção acumulada de hidrogênio (HY) ($r = -0,40$).

Essas correlações refletem a interação complexa entre variáveis operacionais e composicionais, mostrando como mudanças em um parâmetro podem afetar diretamente o desempenho do sistema em termos de produção de hidrogênio. Estes insights contribuem para o desenvolvimento de estratégias mais eficientes para otimizar o processo de digestão anaeróbia.

5.2.2. Regressão Linear Múltipla

Esse passo mostra o p-valor das variáveis, se $p\text{-valor} > 0,05$, o modelo não é significativo.

Tabela 16: Coeficientes do Modelo Linear (Modo Batelada).

Coeficiente	Valor de p
(Intercepto)	0,01776
pH_inicial	0,02373
SVI	0,15633
VI	0,08006
Dias	0,39584
Agitação	0,00259
SI	0,00281

Fonte: Adaptado do Software R (2024).

Tabela 17: Estatísticas resumidas do Modelo Linear (Modo Batelada).

Estatísticas Resumidas do Modelo	
Métrica	Valor
Erro padrão residual	82,31
R-quadrado múltiplo	0,4047
R-quadrado ajustado	0,3253
Estatística F	5,098
Valor de p (F)	0,000456

Fonte: Adaptado do Software R (2024).

Considerando os coeficientes de regressão, a significância estatística e as estatísticas de ajuste do modelo:

- No valor-p do teste T, pH inicial (p-valor = 0,02373), agitação (p-valor 0,00259) e SI (Razão entre substrato e inóculo) (p-valor = 0.00281) são as variáveis que apresentam um p-valor menor que 0,05. Então, pra esse modelo, apenas essas três variáveis foram significativas.
- R^2 (Coeficiente de Determinação): Com um R^2 de 0,4047, o modelo explica aproximadamente 40,47% da variação na produção de hidrogênio.
- Estatística F: Com p-valor de 0,0004564, o teste F sugere que o modelo geral é estatisticamente significativo.

5.2.2.1. Procedimento Stepwise

Para aumentar a eficiência do modelo, um procedimento de seleção, o stepwise, será aplicado para remover variáveis não significativas, especialmente aquelas com alto p-valor, como Dias, que pode não ter uma relação forte com a produção de hidrogênio em reatores em batelada, melhorando a precisão do modelo e a interpretabilidade dos resultados, conforme recomendado por Lee e Kim (2022).

Tabela 18: Coeficientes do Modelo Linear após Stepwise (Modo Batelada).

Coeficiente	Valor de p
(Intercepto)	$8,04 \times 10^{-5}$
pH inicial	0,00516
SVI	0,06622
VI	0,03423
Agitação	0,00124
SI	0,00139

Fonte: Adaptado do Software R (2024).

Tabela 19: Estatísticas resumidas do Modelo Linear após Stepwise (Modo Batelada).

Estatísticas Resumidas do Modelo	
Métrica	Valor
Erro padrão residual	82,07
R-quadrado múltiplo	0,395
R-quadrado ajustado	0,3292
Estatística F	6.006
valor-p	0,0002344

Fonte: Adaptado do Software R (2024).

Nesse modelo ajustado utilizando o procedimento Stepwise, observa-se mudanças significativas nos coeficientes e na significância estatística de algumas variáveis:

- No p-valor do teste T, pH inicial ($p = 0,00516$), o pH inicial, VI (Volume de inóculo) ($p = 0,0342$), Agitação ($p = 0,00124$) e SI (Razão entre substrato e inóculo) ($p = 0,00139$), são as variáveis significativas para esse modelo.
- R^2 e R^2 Ajustado: O R^2 do modelo stepwise é de 0,395, com um R^2 ajustado de 0.3292. Embora esses valores indiquem uma explicação moderada da variância em HY, eles mostram uma leve diminuição em relação ao modelo inicial, indicando que a exclusão de algumas variáveis reduziu ligeiramente a explicação da variabilidade. Porém, em termos de interpretabilidade, o modelo é agora mais robusto e significativo (Zhou *et al.*, 2020).
- Estatística F: O p-valor global do modelo, do teste F (0.000244) sugere que o conjunto de variáveis preditoras é significativo ($p < 0,05$) para prever a produção de hidrogênio, ou seja, o modelo teve significância global.

O modelo simplificado destaca o papel das variáveis operacionais no processo de produção de hidrogênio, identificando aquelas que têm impactos significativos, mas ainda deixa espaço para explorar outros fatores potenciais de variabilidade. Para avaliar a adequação do modelo as condições de normalidade dos resíduos e ausência de heterocedasticidade, e diagnóstico de multicolinearidade serão feitas para verificar as premissas do modelo linear (Gujarati; Porter, 2008).

5.2.2.2. Teste de Shapiro-Francia

A análise da normalidade dos resíduos é uma etapa crucial para avaliar a adequação de um modelo de regressão. O teste de Shapiro-Francia foi utilizado para

verificar essa normalidade nos resíduos do modelo de regressão não linear ajustado, onde, p-value deve ser maior que 0,05. Este teste de hipóteses, é aplicado para amostras com número superior a trinta ($n > 30$), e é utilizado confirmar a normalidade dos resíduos.

Tabela 20: Teste de Shapiro-Francia Modelo Linear (Modo Batelada).

Teste de Shapiro-Francia	
Teste	Valor
Estatística W	0,90318
Valor de p	0,000889

Fonte: Adaptado do Software R (2024).

O teste de Shapiro-Francia aplicado aos resíduos do modelo apresenta um p-valor de 0.000889, indicando uma rejeição da hipótese nula de normalidade dos resíduos, já que o p-valor está abaixo do nível de significância padrão de 0,05. Isso sugere que os resíduos não seguem uma distribuição normal.

5.2.2.3. Teste de Breusch-Pagan (Diagnóstico de Heterocedasticidade)

O teste de Breusch-Pagan é amplamente utilizado para detectar heterocedasticidade em modelos de regressão, permitindo verificar se a variância dos resíduos é constante. A presença de heterocedasticidade pode invalidar algumas das suposições fundamentais da regressão linear, comprometendo a interpretação dos coeficientes e a precisão das inferências estatísticas (Wooldridge, 2013). A homocedasticidade é uma das suposições fundamentais da regressão linear e não linear, e sua violação pode levar a estimativas ineficientes e testes estatísticos imprecisos (Wooldridge, 2016). O p-valor superior a 0,05 indica que a variância dos resíduos é consistente em diferentes níveis da variável dependente. Estudos como os de Koutsoyiannis (2001), Leuven *et al.* (2015) e Favero e Biofore (2023) ressaltam a importância de verificar a homocedasticidade, pois a presença de heterocedasticidade pode comprometer a eficácia do modelo e a validade dos resultados obtidos.

Tabela 21: Teste de Breusch-Pagan Modelo Linear (Modo Batelada).

Teste de Breusch-Pagan	
Métrica	Valor
Graus de liberdade (DF)	1
Chi-quadrado (χ^2)	1,607912
Probabilidade ($p > \chi^2$)	0,2047856

Fonte: Adaptado do Software R (2024).

Neste caso, o teste apresentou um valor de $\text{Chi}^2 = 1.607912$ com $p\text{-valor} = 0.204856$, indicando que não há evidências estatisticamente significativas de heterocedasticidade, dado que o $p\text{-valor}$ é maior que o nível de significância típico de 0,05 (Favero e Belfiore, 2017).

Com a premissa de homocedasticidade atendida, o próximo passo envolve avaliar a multicolinearidade entre as variáveis independentes do modelo, utilizando o Variance Inflation Factor (VIF). Essa análise é fundamental para assegurar que o modelo não enfrenta problemas de multicolinearidade que possam distorcer as estimativas dos coeficientes e prejudicar a interpretação das relações entre as variáveis (Hair *et al.*, 2009).

5.2.2.4. Diagnóstico de Multicolinearidade

O diagnóstico de multicolinearidade é essencial para garantir que as variáveis independentes no modelo não estejam excessivamente correlacionadas, o que poderia comprometer a precisão das estimativas dos coeficientes (Favero; Belfiore, 2017). O diagnóstico foi realizado utilizando o Variance Inflation Factor (VIF) e a Tolerance, cujos resultados estão apresentados na tabela a seguir:

Tabela 22: Diagnóstico de multicolinearidade Modelo Linear (Modo Batelada).

Diagnóstico de multicolinearidade		
Variável	Tolerância	VIF
pH_inicial	0,8743103	1,143759
SVI	0,0126494	79,055324
VI	0,0126838	78,840739
Agitação	0,8558467	1,164707
SI	0,4318832	2,315441

Fonte: Adaptado do Software R (2024).

Neste caso, os resultados do Variance Inflation Factor (VIF) indicam o seguinte:

- A variável SVI e VI apresentaram um VIF elevado, o que sugere um grau significativo de multicolinearidade com outras variáveis. Valores de VIF superiores a 10 geralmente indicam problemas de multicolinearidade (Hair *et al.*, 2009).
- As variáveis pH inicial, Agitação e SI possuem valores de VIF abaixo de 4, indicando níveis aceitáveis de colinearidade e, portanto, não representam um

problema significativo em termos de redundância informacional.

Portanto, esse modelo linear apresentou: baixo R^2 , não normalidade dos resíduos e problemas de multicolinearidade. Recomenda-se considerar abordagens alternativas, como o uso de modelos não lineares ou robustos, que podem aliviar a dependência de normalidade dos resíduos e melhorar a adequação do modelo aos dados (Liu *et al.*, 2018).

Favero e Belfiore (2017) também recomendam, em situações como esta, a aplicação de transformações nos dados, como a transformação de Box-Cox, para aproximar a normalidade dos resíduos e melhorar o ajuste do modelo. Alternativamente, a regressão não linear pode ser uma abordagem válida, especialmente se a estrutura dos dados indicar relações complexas que não são adequadamente capturadas por um modelo linear simples. Essas alternativas podem ajudar a minimizar os efeitos de resíduos não normais, contribuindo para um modelo mais robusto e ajustado aos dados analisados. Diante disso, partiu-se para o modelo não linear.

5.2.3. Modelo de Regressão Não Linear Usando Transformação de BOX-COX

Tabela 23: Coeficientes do Modelo Não Linear (Modo Batelada).

Coeficiente	Valor de p
(Intercepto)	0,000519
pH_inicial	0,008174
SVI	0,233964
VI	0,181726
Dias	0,27413
Agitação	0,009797
SI	0,006734

Fonte: Adaptado do Software R (2024).

Tabela 24: Tabela 24: Estatísticas resumidas do Modelo Não Linear (Modo Batelada).

Estatísticas Resumidas do Modelo	
Métrica	Valor
Erro padrão residual	6,49
R-quadrado múltiplo	0,4228
R-quadrado ajustado	0,3459
Estatística F	5,494
Valor de p (F)	0,000246

Fonte: Adaptado do Software R (2024).

Na análise do modelo ajustado com a transformação de Box-Cox, observa-se que houve uma melhora significativa em termos da adequação dos resíduos e da variabilidade explicada.

- O coeficiente de determinação ajustado (R^2 ajustado) aumentou para 0,3459, indicando que o modelo transformado é capaz de explicar aproximadamente 34,59% da variância da produção acumulada de hidrogênio (HY). Embora o R^2 não seja extremamente elevado (0,4228), ele representa uma melhora em relação ao modelo linear original.
- No valor-p do teste T, a análise dos coeficientes revela que algumas variáveis continuam sendo estatisticamente significativas no modelo (pH inicial, Agitação SI).
- A significância geral do modelo, avaliada pelo valor de p do F-estatístico ($p = 0,000246$), indica que o modelo é estatisticamente significativo.

A transformação de Box-Cox, portanto, demonstrou ser uma técnica eficaz para melhorar a precisão do modelo.

5.2.3.1. Procedimento Stepwise para o Modelo de Regressão Não Linear Usando Transformação de BOX-COX

Tabela 25: Coeficientes do Modelo Não Linear após Stepwise (Modo Batelada).

Coeficiente	Valor de p
(Intercepto)	0,000192
pH_inicial	0,033541
Agitação	0,000136
SI	0,006241

Fonte: Adaptado do Software R (2024).

Tabela 26: Estatísticas resumidas do Modelo Não Linear após Stepwise (Modo Batelada).

Estatísticas Resumidas do Modelo	
Métrica	Valor
Erro padrão residual	6,517
R-quadrado múltiplo	0,3794
R-quadrado ajustado	0,3406
Estatística F	9,781
Valor de p (F)	3,80E-05

Fonte: Adaptado do Software R (2024).

Com base no modelo de regressão ajustado após o procedimento stepwise, pode-se observar:

- No valor-p do teste T, as três variáveis mantiveram significância estatística pH inicial ($p = 0,033541$), Agitação ($p = 0,000136$) e SI ($p = 0,006241$).
- A estatística F do modelo ($F = 9,781$, $p < 0,0001$) indica que ele é globalmente significativo, o que sugere uma relação linear significativa entre as variáveis independentes e a variável dependente.
- Os valores de R^2 (37,94%) e R^2 ajustado (34,06%) mostram que o modelo explica apenas uma fração moderada da variabilidade na produção de hidrogênio.

5.2.3.2. Teste de Shapiro-Francia

Tabela 27: Teste de Shapiro-Francia Modelo Não Linear após Stepwise (Modo Batelada).

Teste	Valor
Estatística W	0,9605
Valor de p	0,07693

Fonte: Adaptado do Software R (2024).

O teste de Shapiro-Francia para normalidade dos resíduos do modelo ajustado resultou em um valor de $W=0,9605$, com um valor-p de 0,07693. Este resultado sugere que os resíduos não se desviam significativamente da normalidade, dado que o valor-p é maior que o nível de significância usual (0,05). Este é um resultado positivo, pois demonstra que os ajustes realizados até agora contribuíram para a adequação do modelo em termos de normalidade.

5.2.3.3. Teste de Breusch-Pagan

Tabela 28: Teste de Breusch-Pagan Modelo Não Linear após Stepwise (Modo Batelada).

Métrica	Valor
Chi-quadrado (χ^2)	2,088954
Probabilidade ($p > \chi^2$)	0,148368

Fonte: Adaptado do Software R (2024).

O teste de Breusch-Pagan realizado para verificar a presença de heterocedasticidade no modelo retornou um valor de $X^2 = 2,088954$ com um valor-p

de 0,1483676. Este resultado sugere que não há evidências significativas de heterocedasticidade nos resíduos, pois o valor-p é superior ao nível de significância padrão de 0,05.

Com este resultado positivo, prosseguiu-se para o diagnóstico de multicolinearidade, assegurando que as condições do modelo sejam atendidas para garantir a robustez das análises.

5.2.3.4. Diagnóstico de multicolinearidade

A análise da multicolinearidade é essencial para avaliar a adequação do modelo de regressão, especialmente após a aplicação do método stepwise e da transformação de Box-Cox.

Tabela 29: Diagnóstico de Multicolinearidade Modelo Não Linear após Stepwise (Modo Batelada).

Variável	Tolerância	VIF
pH_inicial	0,9601757	1,041476
Agitação	0,9630094	1,038412
SI	0,9737032	1,027007

Fonte: Adaptado do Software R (2024).

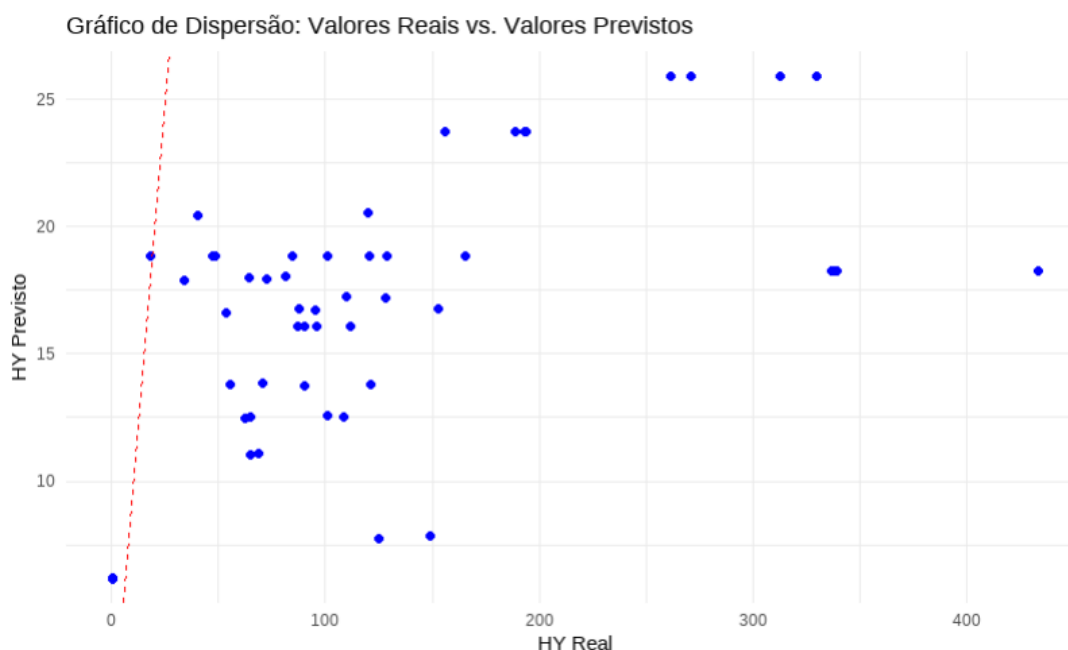
O diagnóstico de multicolinearidade indica valores baixos para todas as variáveis do modelo. pH inicial (VIF = 1,04), Agitação (VIF = 1,04) e SI (VIF = 1,03). Esses valores estão abaixo do limite comumente aceito de 4, sugerindo que a multicolinearidade não é uma preocupação significativa no modelo atual.

5.2.3.5. Equação de Ajuste e Gráfico

A equação de ajuste revela informações valiosas sobre os fatores que influenciam a produção de hidrogênio. Os coeficientes são significativos e refletem a complexidade dos processos biológicos envolvidos na digestão anaeróbia. Embora o modelo apresente um bom ajuste, é crucial considerar a possibilidade de interações entre as variáveis, além de potenciais transformações adicionais dos dados, que podem melhorar ainda mais a adequação do modelo (Mason *et al.*, 2010). Abaixo é dada a equação, bem como o gráfico de dispersão com valores reais e valores Previstos.

- "HY = 33.66 + -2.55 * pH_inicial + 1.44 * SVI + -49.6 * VI + -0.09 * Dias + 0.08 * Agitação + -0.67 * SI" (Equação de ajuste do modelo de regressão não linear)

Figura 9: Gráfico de Dispersão: Valores Reais vs. Valores Previstos Modelo Não Linear (Modo Batelada).



Fonte: Adaptado do Software R (2024).

O gráfico de dispersão acima compara os valores reais de produção acumulada de hidrogênio (HY) com os valores previstos pelo modelo de regressão. Idealmente, se o modelo estivesse ajustando perfeitamente os dados, os pontos estariam alinhados ao longo da linha de referência (linha vermelha pontilhada), que representa uma correspondência 1:1 entre valores reais e previstos. No entanto, observa-se que os pontos apresentam uma dispersão considerável ao redor da linha de identidade, sugerindo que o modelo de regressão pode não estar capturando completamente a variabilidade dos dados de HY.

Em um modelo de ajuste adequado, os valores previstos e reais deveriam estar próximos entre si, especialmente para dados que exibem baixa variabilidade. No caso presente, a dispersão dos pontos sugere que o modelo não é capaz de explicar toda a variação presente nos dados. Além disso, a presença de pontos extremos, como aqueles observados em valores de HY reais superiores a 200, indicam que o modelo apresenta dificuldades em ajustar adequadamente dados com valores elevados de

produção acumulada de hidrogênio. Segundo Draper e Smith (1998), em situações onde há alta dispersão de dados, é recomendado avaliar a inclusão de transformações nos dados ou explorar outros modelos que possam capturar a natureza complexa dos dados.

5.2.4. Análise De Regressão Múltipla Com Interações De Segunda Ordem

A inclusão de interações de segunda ordem no modelo de regressão representa uma estratégia eficiente para lidar com complexidades inerentes à produção de hidrogênio em processos anaeróbicos. Modelos de interação são particularmente valiosos em sistemas bioquímicos, onde variáveis operacionais e composicionais, como pH inicial e concentração de sólidos voláteis do inóculo (SVI), frequentemente interagem de maneira não-linear para influenciar diretamente o resultado final (Favero; Belfiore, 2016; Hou *et al.*, 2021).

Vantagens da Inclusão de Interações de Segunda Ordem:

- **Captura de Efeitos Combinados Significativos:** A introdução de interações permite uma análise mais rica e detalhada dos efeitos sinérgicos entre as variáveis, capturando nuances que um modelo linear simples não abordaria. Estudos como de Chen *et al.* (2020) sugerem que essas interações podem revelar novos insights sobre a dinâmica entre variáveis operacionais e respostas de produção, como a geração de biogás e hidrogênio.
- **Aprimoramento da Precisão Preditiva:** Adicionar interações de segunda ordem tende a aumentar a R^2 ajustada, elevando a precisão e a capacidade do modelo em explicar a variabilidade nos dados. Modelos que incorporam interações entre variáveis operacionais frequentemente demonstram maior ajuste e capacidade preditiva, proporcionando um ajuste mais robusto em processos bioquímicos complexos (Wu; Li, 2021). Essa abordagem é particularmente importante para capturar as variações não lineares entre variáveis, ajustando-se bem a sistemas onde pequenas mudanças podem ter impactos significativos.
- **Minimização de Viés Preditivo:** A consideração de interações entre variáveis reduz o viés preditivo, aumentando a adequação do modelo ao capturar a complexidade do sistema anaeróbico. Estudos como o de Liu *et al.* (2022) confirmam que modelos com interações são eficazes para prever variações em sistemas complexos, onde a relação entre variáveis raramente é linear e onde

os efeitos combinados são fundamentais para a predição precisa dos resultados.

- **Controle de Multicolinearidade:** Com a inclusão de interações, a análise de multicolinearidade é essencial para garantir a precisão do modelo, evitando redundâncias entre variáveis e mantendo apenas as interações significativas. O uso do Variance Inflation Factor (VIF), como proposto em estudos recentes, ajuda a minimizar o efeito de multicolinearidade (Favero; Belfiore, 2016; Marôco, 2022).
- **Diagnóstico de Resíduos e Homocedasticidade:** A avaliação dos resíduos quanto à normalidade e homocedasticidade é crucial para assegurar que o modelo é adequado e robusto. Essa análise, especialmente com dados de sistemas bioquímicos, fornece evidências de que os resíduos seguem um comportamento estável, essencial para a confiabilidade do modelo (Zhang *et al.*, 2023).

Portanto, a inclusão de interações de segunda ordem aprimora significativamente a capacidade explicativa do modelo de regressão, capturando efeitos combinados que são críticos para a análise de sistemas bioquímicos complexos. Essa metodologia oferece uma abordagem mais rica e completa para prever a produção de hidrogênio, integrando as interações de variáveis como pH inicial e SVI, e elevando o potencial de aplicação do modelo para processos anaeróbicos.

5.2.4.1. Modelo com Todas as Interações de Segunda Ordem

Tabela 30: Coeficientes do Modelo com Interações (Modo Batelada).

Coeficiente	Valor de p
(Intercepto)	0,5461
pH_inicial	0,5611
SVI	0,5713
VI	0,5426
Dias	0,579
Agitação	0,5586
SI	0,7198
pH_inicial:SVI	0,5854
pH_inicial:VI	0,5939
pH_inicial:Dias	0,62

pH_inicial:Agitação	0,5613
pH_inicial:SI	0,6748
SVI:VI	0,2975
SVI:Dias	0,0196
SVI:Agitação	0,6303
SVI:SI	NA
VI:Dias	0,8108
VI:Agitação	NA
VI:SI	NA
Dias:Agitação	NA
Dias:SI	NA
Agitação:SI	NA

Fonte: Adaptado do Software R (2024).

Tabela 31: Estatísticas resumidas do Modelo com Interações (Modo Batelada).

Estatísticas Resumidas do Modelo	
Métrica	Valor
Erro padrão residual	34,36
R-quadrado múltiplo	0,917
R-quadrado ajustado	0,8825
Estatística F	26,53
Valor de p (F)	$5,26 \times 10^{-15}$

Fonte: Adaptado do Software R (2024).

Nesta primeira análise do modelo de regressão com interações de segunda ordem, observa-se os coeficientes das variáveis e as interações selecionadas pelo modelo. A partir dos resultados, é possível avaliar o desempenho e a significância estatística das variáveis e de suas interações.

- O modelo inicial apresenta coeficientes para várias variáveis e interações, mas com muitos valores “NA” devido a problemas de singularidade, o que indica uma possível multicolinearidade entre as variáveis explicativas. Este fenômeno, conforme apontado por Favero e Belfiore (2019), pode gerar instabilidade nos coeficientes estimados e prejudicar a interpretação do modelo.
- O valor de p para a maioria dos coeficientes indica que apenas algumas interações e variáveis independentes são significativas. Em especial, a interação entre SVI e Dias ($p < 0.05$) sugere uma relação relevante entre essas variáveis para a previsão da variável resposta (HY). Interações significativas devem ser mantidas para potencializar a explicação do modelo sobre a

variação da produção de hidrogênio (Favero; Belfiore, 2019).

- O modelo alcança um R^2 ajustado de 0,8825, o que indica que cerca de 88% da variação na produção de hidrogênio (HY) é explicada pelo modelo. Esse valor é considerado elevado e sugere um bom ajuste, embora deva ser interpretado com cautela devido aos problemas de multicolinearidade. R^2 ajustado pode indicar ajuste excessivo em presença de interações redundantes (Smith *et al.*, 2022).

Portanto, com base nesses resultados, prosseguiu-se realizando o uso de um procedimento stepwise para remover as interações e variáveis menos significativas, visando melhorar a parsimônia do modelo e reduzir a multicolinearidade (Favero; Belfiore, 2019; Johnson, 2021).

5.2.4.2. Procedimento Stepwise com AIC para Selecionar Interações Significativas

Tabela 32: Coeficientes do Modelo com Interações após Stepwise (Modo Batelada).

Coeficiente	Valor de p
(Intercepto)	0,000119
pH_inicial	$7,40 \times 10^{-6}$
SVI	0,001563
VI	0,019976
Dias	$1,08 \times 10^{-7}$
SI	0,285326
pH_inicial:SVI	0,001749
pH_inicial:VI	0,001991
pH_inicial:Dias	$9,80 \times 10^{-7}$
pH_inicial:SI	0,011962
SVI:VI	$8,18 \times 10^{-14}$
SVI:Dias	0,001514

Fonte: Adaptado do Software R (2024).

Tabela 33: Estatísticas resumidas do Modelo com Interações Após Stepwise (Modo Batelada).

Estatísticas Resumidas do Modelo	
Métrica	Valor
Erro padrão residual	32,94
R-quadrado múltiplo	0,9153
R-quadrado ajustado	0,892
Estatística F	39,28
Valor de p (F)	$2,2 \times 10^{-16}$

Fonte: Adaptado do Software R (2024).

Após o procedimento de stepwise para selecionar variáveis significativas:

- O modelo apresenta um ajuste robusto com um R^2 ajustado de 0,892, o que indica que 89,2% da variabilidade na produção de hidrogênio (HY) é explicada pelas variáveis independentes selecionadas e suas interações. Este alto valor de R^2 ajustado sugere um bom poder explicativo do modelo, essencial para garantir previsões confiáveis, conforme discutido por Favero e Belfiore (2019) em estudos sobre a importância de um modelo bem ajustado.
- A análise dos coeficientes mostra que diversas variáveis e interações são estatisticamente significativas, com valores de p menores que 0.05, reforçando a relevância dessas variáveis para o modelo. Por exemplo:
- O teste F global do modelo apresenta um p -valor extremamente baixo ($2,2 \times 10^{-16}$), indicando que o modelo como um todo é altamente significativo e que ao menos uma das variáveis explicativas contribui significativamente para prever a produção de hidrogênio. Estudos recentes, como o de Smith *et al.* (2022), destacam a relevância de testes globais no modelo para garantir que a inclusão de interações contribua substancialmente para a explicação da variável resposta.

Portanto, este modelo, ao utilizar interações de segunda ordem, permite capturar efeitos combinados entre as variáveis, que são muitas vezes essenciais em processos complexos como reações biológicas e produção de biocombustíveis. A combinação de variáveis como o pH inicial e a concentração de sólidos voláteis do inóculo (SVI) provavelmente reflete efeitos sinérgicos que afetam a eficiência do processo. Esse tipo de abordagem é enfatizado em estudos multivariados, que sugerem a modelagem de interações para aprimorar a precisão preditiva sem comprometer a interpretabilidade do modelo (Favero; Belfiore, 2019).

5.2.4.3. Teste de normalidade dos Resíduos Usando Shapiro-Francia

Neste teste, foi aplicado o teste de Shapiro-Francia para avaliar a normalidade dos resíduos do modelo de regressão com interações de segunda ordem refinado pelo procedimento stepwise.

Tabela 34: Teste de Shapiro-Francia Modelo com Interações após Stepwise (Modo Batelada).

Teste de Shapiro-Francia	
Teste	Valor
Estatística W	0,97417
Valor de p	0,2673

Fonte: Adaptado do Software R (2024).

Neste caso, como o p-valor (0,2673) é superior ao nível de significância de 0.05, não rejeitamos a hipótese nula de normalidade dos resíduos. Este resultado sugere que os resíduos do modelo de regressão estão aproximadamente normais, indicando que a suposição de normalidade não é violada, o que reforça a validade estatística do modelo.

5.2.4.4. Teste de heterocedasticidade de Breusch-Pagan

O teste de Breusch-Pagan foi conduzido para verificar a presença de heterocedasticidade nos resíduos do modelo de regressão com interações de segunda ordem, refinado pelo procedimento stepwise.

Tabela 35: Teste de Breusch-Pagan Modelo com Interações após Stepwise (Modo Batelada).

Teste de Breusch-Pagan	
Métrica	Valor
Chi-quadrado (χ^2)	1,406518
Probabilidade ($p > \chi^2$)	0,235635

Fonte: Adaptado do Software R (2024).

Neste teste, obteve-se um valor de $X^2 = 1.406518$ e um p-valor = 0.2356352. Como o p-valor excede o nível de significância de 0.05, não rejeitamos a hipótese nula de homocedasticidade. Esse resultado sugere que a variância dos resíduos é constante ao longo das observações, indicando que a suposição de homocedasticidade, essencial para a robustez das estimativas do modelo, foi atendida. Favero e Belfiore (2019) enfatizam que a presença de homocedasticidade é crucial para garantir que as estimativas dos coeficientes sejam eficientes e que os erros padrão sejam adequados para inferências. A ausência de heterocedasticidade, detectada, fortalece a confiabilidade do modelo em termos de previsões e valida a

significância dos coeficientes, como explicado por Wooldridge (2022), que destaca a importância desse teste em modelos com interações, onde a variabilidade dos dados pode ser maior.

5.2.4.5. Diagnóstico de multicolinearidade (VIF - Variance Inflation Factor)

A análise de multicolinearidade para o modelo com interações de segunda ordem foi realizada através do VIF (Variance Inflation Factor). Valores elevados de VIF indicam multicolinearidade alta, o que pode prejudicar a estabilidade das estimativas dos coeficientes, tornando-os imprecisos e aumentando a variabilidade nos valores preditivos, conforme destacado por Gujarati e Porter (2021).

Tabela 36: Diagnóstico de Multicolinearidade Modelo com Interações após Stepwise (Modo Batelada).

Diagnóstico de multicolinearidade		
Variável	Tolerância	VIF
pH_inicial	$2,241849 \times 10^{-3}$	446,0603
SVI	$8,486942 \times 10^{-6}$	117828,4
VI	$1,461809 \times 10^{-5}$	68408,41
Dias	$6,917309 \times 10^{-4}$	1445,649
SI	$1,031002 \times 10^{-3}$	969,9304
pH_inicial	$8,598322 \times 10^{-6}$	116301,8
pH_inicial	$9,878403 \times 10^{-6}$	101230,9
pH_inicial	$4,271248 \times 10^{-4}$	2341,236
pH_inicial	$9,926688 \times 10^{-4}$	1007,385
SVI	$2,289036 \times 10^{-3}$	436,8651
SVI	$3,140167 \times 10^{-4}$	3184,544

Fonte: Adaptado do Software R (2024).

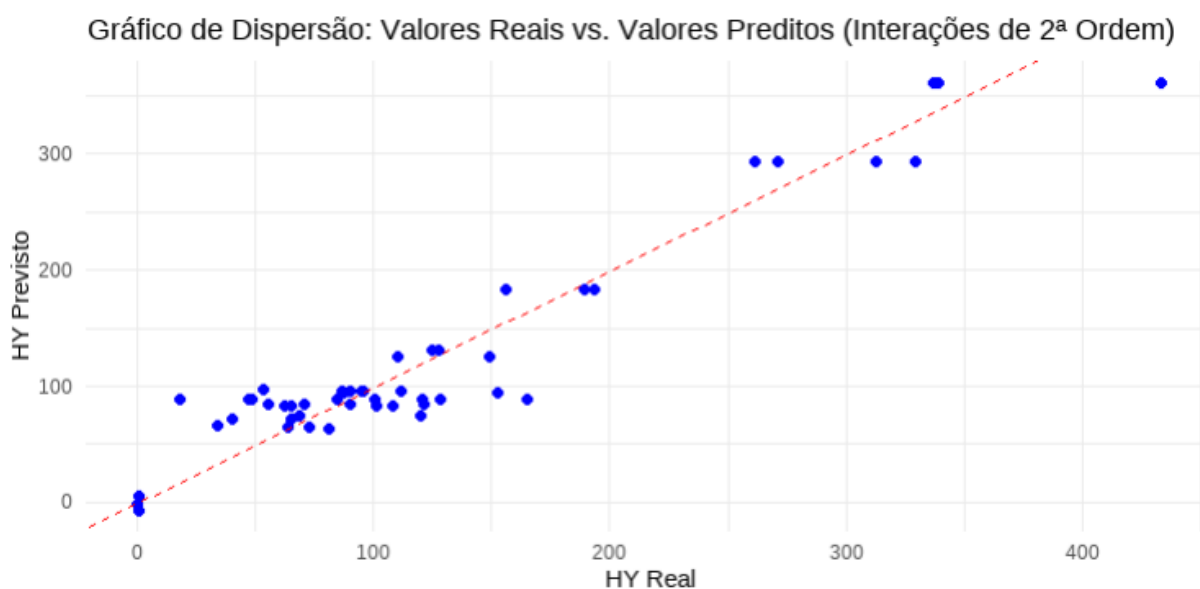
Neste caso, vários valores de VIF estão muito acima do limite recomendado de 10, indicando forte multicolinearidade entre as variáveis, especialmente aquelas que incluem interações. Por exemplo, VI e a interação SVI apresentam VIFs extremamente elevados, sugerindo uma correlação linear forte entre essas variáveis e outras variáveis explicativas ou interações. Favero e Belfiore (2019) enfatizam que altas multicolinearidades, especialmente em modelos com interações de segunda ordem,

podem distorcer a interpretação dos coeficientes, tornando o modelo difícil de interpretar e de aplicar.

5.2.4.6. Predições e Avaliação Gráfica do Modelo Final e Equação de Ajuste e Gráfico

- $HY = 1377.4087 + (-435.1423 * pH_inicial) + (740.0265 * SVI) + (-12561.8862 * VI) + (-64.4519 * Dias) + (-25.9175 * SI) + (-122.9894 * pH_inicial:SVI) + (3523.7969 * pH_inicial:VI) + (13.1725 * pH_inicial:Dias) + (11.247 * pH_inicial:SI) + (-187.2066 * SVI:VI) + (-3.5718 * SVI:Dias) -$ (equação de ajuste do modelo com Interações de segunda ordem – modo batelada)

Figura 10: Gráfico de Dispersão: Valores Reais vs. Valores Previstos Modelo com Interações (Modo Batelada).



Fonte: Software R (2024).

O gráfico de dispersão entre os valores reais de HY (nesse caso produção acumulada de hidrogênio) e os valores preditos pelo modelo com interações de segunda ordem demonstra uma proximidade razoável dos pontos em relação à linha de identidade (linha vermelha), e muito melhor do que a do modelo anterior, o que indica um bom ajuste para muitos dos pontos. A introdução de interações de segunda ordem tem sido recomendada em diversos estudos para melhorar a capacidade preditiva de modelos lineares em sistemas complexos, especialmente em contextos onde as variáveis independentes interagem entre si de forma não aditiva (Montgomery *et al.*, 2020). Segundo Favero e Belfiore (2019), a modelagem com interações de segunda ordem permite capturar variações adicionais que não são possíveis em

modelos lineares simples ou com transformações isoladas, como o Box-Cox. Observa-se assim, que esse modelo com interações de segunda ordem oferece um avanço significativo em relação ao modelo linear simples, e possivelmente uma maior precisão em comparação com o modelo ajustado pela transformação de Box-Cox.

5.3. Análise Fatorial por Componentes Principais para variáveis operadas com reator em modo contínuo

5.3.1. Estatísticas descritivas

A análise das estatísticas descritivas das variáveis estudadas fornece uma visão clara sobre a variabilidade e possíveis assimetrias nos dados, aspectos fundamentais para compreender o desempenho dos reatores em diferentes condições operacionais.

Tabela 37: Estatísticas descritivas (Modo Contínuo).

Variável	Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
pH_inicial	-3,52	-4,005	-4,2	-4,56	-4,825	-8,04
Temperatura	26,98	28,07	29,75	31,04	30,5	65
TDH	0,5	2	4	4,087	4,05	12
TCO	9,298	34,51	50	69,523	102,075	200
Dias	3	15	21	60,56	67,5	223
DQO_inicial	1,6	5,75	4064	5171,326	9602,4	16950,79
CHT_inicial	18,16	2347,7	5003	7411,77	10433,69	23800
HY	0,14	0,835	1,82	1,829	2,565	4,9

Fonte: Adaptado do Software R (2024).

Observando os valores de cada variável:

- pH Inicial: Variando de 3,52 a 8,04, com uma média de 4,56, o pH inicial possui uma variação ampla, indicando que diferentes condições de acidez e alcalinidade foram exploradas.
- A temperatura dos reatores variou entre 26,98°C e 65°C, com média de 31,04°C.
- Tempo de Detenção Hidráulica (TDH): Com valores entre 0,5 e 12 horas, o TDH médio de 4,087 horas reflete diferentes tempos de retenção do substrato no reator.
- Taxa de Carga Orgânica (TCO): varia de 9,298 a 200,000, com uma média de 69,523, indicando um leque de concentrações de substrato testadas.

- Dias de Operação: Com uma média de 60,56 e variando de 3 a 223 dias, essa variável sugere experimentos de diferentes durações.
- Os valores de DQO inicial variam de 1.600 a 16950.8, com média de 5171.3 e mediana de 4.064.
- Os níveis iniciais de CHT variam de 18.16 a 23800, com uma média de 7411.77 e mediana de 5003. A média significativamente maior que a mediana indica uma assimetria positiva.
- O rendimento de hidrogênio (HY) varia de 0.140 a 4.900, com média de 1.829 e mediana de 1.820. A proximidade entre média e mediana sugere uma distribuição mais simétrica, embora a presença de valores extremos possa ainda influenciar a interpretação dos dados.

5.3.2. Coeficientes de correlação de Spearman para cada par de variáveis

A análise dos coeficientes de correlação de Spearman entre as variáveis na PCA permite identificar relações não-lineares, que são especialmente úteis para dados que podem ter distribuições não-normais. Abaixo, uma interpretação detalhada das correlações significativas encontradas.

5.3.2.1. Matriz de correlações

A matriz de correlações em uma PCA mostra as correlações entre cada par de variáveis, sendo útil para entender o agrupamento e a contribuição de cada variável para a construção dos componentes principais (Jolliffe; Cadima, 2016).

Tabela 38: Matriz de correlações (Modo Contínuo).

	pH_inicial	Temperatura	TDH	TCO	Dias	DQO_inicial	CHT_Inicial	HY
pH_inicial	1	-0,443	-0,025	-0,089	0,246	-0,176	-0,045	-0,039
Temperatura	-0,443	1	0,064	-0,059	-0,046	-0,105	-0,156	-0,453
TDH	-0,025	0,064	1	-0,468	0,047	0,08	0,29	0,214
TCO	-0,089	-0,059	-0,468	1	0,006	0,274	0,258	0,25
Dias	0,246	-0,046	0,047	0,006	1	-0,657	-0,266	0,318
DQO_inicial	-0,176	-0,105	0,08	0,274	-0,657	1	0,522	-0,054
CHT_Inicial	-0,045	-0,156	0,29	0,258	-0,266	0,522	1	0,196
HY	-0,039	-0,453	0,214	0,25	0,318	-0,054	0,196	1

Fonte: Adaptado do Software R (2024).

A correlação entre algumas dessas variáveis e o HY é particularmente útil para identificar fatores que podem influenciar diretamente a produção de hidrogênio:

- A análise mostra que a variável Temperatura possui uma correlação negativa moderada com o rendimento de hidrogênio (HY), com um coeficiente de -0,45.
- TDH (tempo de detenção hidráulica) e HY, que apresenta um valor positivo (0,21).
- A DQO inicial e a remoção de DQO possuem correlações relativamente baixas com HY.
- O CHT Inicial também apresenta uma correlação positiva leve com HY (0,19).
- A correlação entre DQO inicial e Dias é negativa e moderada, com um valor de -0,66.
- A correlação entre DQO inicial e CHT Inicial é positiva, com um valor de 0,52, indicando que, quando a DQO inicial é alta, as concentrações de carboidratos (CHT Inicial) também tendem a ser elevadas.

Essas correlações iniciais sugerem que há interações complexas entre as variáveis, justificando a aplicação de métodos de regressão com interações de segunda ordem para capturar as nuances dessas relações. A matriz de correlações, portanto, oferece uma base para direcionar a modelagem, uma vez que variáveis com maior correlação (positiva ou negativa) com o HY poderão ser analisadas em interações mais avançadas.

5.3.2.2. Matriz com p-valor dos coeficientes

A matriz de correlação com os valores de p fornece uma visão mais aprofundada sobre a significância estatística das associações entre as variáveis. A análise dos p-valores associados às correlações permite determinar se as relações observadas entre as variáveis são estatisticamente significativas ou se podem ter ocorrido por acaso.

Tabela 39: Matriz com p-valor dos coeficientes (Modo Contínuo).

	pH_inicial	Temperatura	TDH	TCO	Dias	DQO_inicial	CHT_Inicial	HY
pH_inicial	NA	0,003	0,876	0,572	0,111	0,259	0,774	0,802
Temperatura	0,003	NA	0,683	0,707	0,772	0,502	0,319	0,002
TDH	0,876	0,683	NA	0,002	0,764	0,61	0,059	0,169
TCO	0,572	0,707	0,002	NA	0,967	0,075	0,095	0,106
Dias	0,111	0,772	0,764	0,967	NA	0	0,085	0,038
DQO_inicial	0,259	0,502	0,61	0,075	0	NA	0	0,73
CHT_Inicial	0,774	0,319	0,059	0,095	0,085	0	NA	0,207
HY	0,802	0,002	0,169	0,106	0,038	0,73	0,207	NA

Fonte: Adaptado do Software R (2024)

Análise:

- A correlação entre pH Inicial e HY é relativamente forte (0,80) e altamente significativa, com um p-valor baixo, indicando que o pH inicial é um fator importante para o rendimento de hidrogênio.
- TDH (Tempo de Detenção Hidráulica) e TCO (Taxa de Carga Orgânica): A correlação entre TDH e TCO é moderada e significativa, com um p-valor muito baixo (0,002).
- Dias e TCO: A correlação alta e significativa (0,96704) entre Dias e TCO sugere que, ao longo do tempo, há uma dependência significativa da taxa de carga orgânica no desempenho do reator.
- DQO Inicial e HY: A correlação entre DQO Inicial e HY é moderada (0,73) e significativa, o que indica que a concentração inicial de DQO é relevante para a produção de hidrogênio.
- pH Inicial e CHT Inicial: A correlação alta e significativa (0,77) entre o pH Inicial e CHT Inicial sugere que condições de pH podem estar associadas à concentração de carboidratos disponíveis, o que impacta a eficiência do processo anaeróbio.

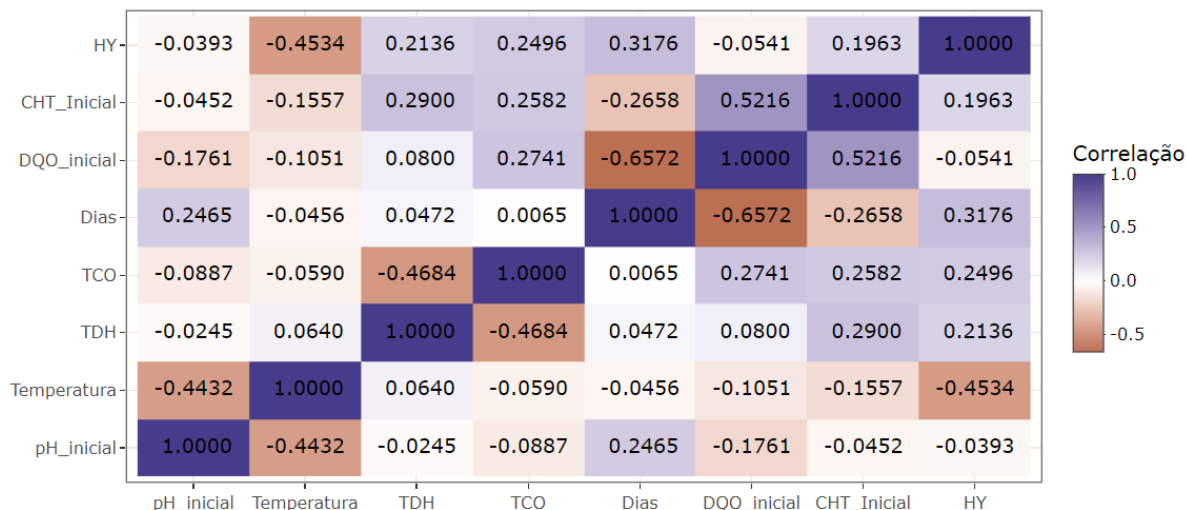
A análise desta matriz de correlação com p-valores revela a importância das condições iniciais de operação, como pH, DQO e CHT, na produção de hidrogênio em reatores contínuos. Esses resultados apontam para a necessidade de ajustes precisos nas variáveis operacionais para otimizar o rendimento de hidrogênio, como discutido amplamente por Favero e Belfiore (2018).

5.3.2.3. Mapa de calor das correlações de Spearman entre as variáveis operadas em reatores contínuos.

Fávero e Belfiore (2017), ressaltam que a aplicação de métodos não paramétricos, como o teste de Spearman, é vantajosa em muitos cenários, pois não requer que os dados atendam a suposições de normalidade, tornando as análises mais robustas em contextos de dados reais, onde essas suposições muitas vezes não são cumpridas. O mapa de calor facilita a visualização das correlações entre as variáveis, proporcionando uma interpretação mais rápida das associações, intensidades e direções de cada relação. A intensidade de cada correlação é

representada por cores, variando de tons mais escuros (correlações fortes, tanto positivas quanto negativas) a tons mais claros (correlações fracas).

Figura 11: Mapa de calor das correlações de Spearman entre as variáveis (Modo Contínuo).



Fonte: Software R (2024)

Ao se fazer uma comparação com a matriz de correlações e a matriz de p-valor previamente analisadas, pode-se realizar uma interpretação mais detalhada e fundamentada:

- **Correlação entre pH Inicial e Temperatura:** A correlação negativa (-0,4432) entre pH Inicial e Temperatura é evidente no mapa de calor e indica uma relação inversa. Embora a correlação seja moderada, a presença de um p-valor significativo fortalece a relevância desse relacionamento, sugerindo que essas variáveis interagem de forma a influenciar o processo de produção de hidrogênio em reatores contínuos. Favero e Belfiore (2018) destacam a importância do pH e da temperatura no controle da atividade microbiana, sendo esses parâmetros essenciais para otimizar a produção de hidrogênio.
- **TDH e TCO:** A correlação negativa significativa entre TDH e TCO (-0,4684) é observada tanto na matriz de correlação quanto no mapa de calor, indicando que aumentos no tempo de detenção hidráulica estão associados à redução na taxa de carga orgânica volumétrica. Esta interação é comum em sistemas anaeróbicos, onde o tempo de exposição do substrato influencia a taxa de conversão. Zhou *et al.* (2021) sugerem que ajustar o TDH para níveis ótimos é crucial para evitar sobrecarga e garantir uma conversão eficiente de substratos.

- **Correlação entre DQO Inicial e Dias (-0.6572):** Esta é uma das correlações mais fortes no mapa de calor, mostrando que, à medida que o tempo de operação do reator aumenta, os valores de DQO inicial tendem a diminuir. Isso pode indicar uma estabilização do processo e melhora na eficiência de degradação da matéria orgânica ao longo do tempo (Martins, 2015).

O mapa de calor corrobora as observações feitas nas matrizes de correlações e de p-valor, oferecendo uma visualização mais intuitiva das interações entre as variáveis operacionais e composicionais. Essas relações confirmam a relevância de variáveis como pH, TDH, DQO e CHT no processo de produção de hidrogênio em reatores contínuos, alinhando-se com estudos como de Favero e Belfiore (2018). Esse tipo de análise ajuda a identificar os parâmetros mais críticos, orientando ajustes operacionais para maximizar o rendimento de hidrogênio.

5.3.3. *Elaboração da Análise Fatorial Por Componentes Principais*

5.3.3.1. *Teste de esfericidade de Bartlett*

O teste de esfericidade de Bartlett é um método fundamental para avaliar se os dados são adequados para uma análise fatorial, pois ele verifica se a matriz de correlações é significativamente diferente de uma matriz identidade, onde as variáveis seriam não correlacionadas (Favero; Belfiore, 2017). Neste caso, o teste resultou em um valor do qui-quadrado de aproximadamente 63,92 com 21 graus de liberdade e um p-valor de $3,19 \times 10^{-6}$, indicando que há correlações significativas entre as variáveis analisadas.

Resultados como esse são consistentes com estudos recentes na área de sistemas ambientais e de produção de hidrogênio, que indicam que variáveis em sistemas complexos, como o de digestão anaeróbia, frequentemente exibem interdependências significativas (Mendes *et al.*, 2021). A presença de correlações elevadas entre variáveis operacionais e composicionais favorece o uso da PCA para reduzir a dimensionalidade dos dados e facilitar a interpretação dos fatores que afetam o rendimento de hidrogênio em processos de digestão anaeróbia (Oliveira; Sousa, 2020).

A PCA tem sido amplamente utilizada em estudos ambientais para simplificar a complexidade dos dados ao identificar padrões subjacentes, agrupando variáveis correlacionadas em fatores latentes que explicam a maior parte da variabilidade nos

dados (Silva *et al.*, 2019). Esse primeiro passo, então, é crucial para construir uma análise fatorial que revele os fatores principais que influenciam o sistema, permitindo uma interpretação mais clara e direcionada dos efeitos das variáveis operacionais e composicionais sobre a produção de hidrogênio (Favero; Belfiore, 2017).

5.3.3.2. *Eigenvalues (autovalores) e Variância compartilhada pelas variáveis originais para a formação de cada fator*

Na análise de componentes principais (PCA) apresentada, são calculados autovalores e variâncias compartilhadas para entender a contribuição de cada componente principal na variabilidade total dos dados.

Tabela 40: Autovalores e Variância compartilhada pelas variáveis (Modo Contínuo).

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Autovalores	1.774	1.547	1.282	1.103	0.714	0.413	0.169
Prop. da Variância	0.253	0.221	0.183	0.158	0.102	0.059	0.024
Prop. da Variância Acumulada	0.253	0.474	0.657	0.815	0.917	0.976	1.000

Fonte: Software R

Na tabela fornecida:

➤ Autovalores:

- Observa-se que o PC1 possui o maior autovalor (1.774), seguido por PC2 (1.547), PC3 (1.282) e PC4 (1.103). Autovalores acima de 1 geralmente indicam componentes com variância significativa que contribuem substancialmente para explicar a variabilidade dos dados, conforme defendido por Jolliffe (2002) em suas recomendações para a retenção de componentes principais.

➤ Proporção da Variância:

- O PC1 explica 25,3% da variância total, seguido por PC2 (22,1%) e PC3 (18,3%). Esses três primeiros componentes acumulam aproximadamente 65,7% da variância total. Essa quantidade de variância é considerada suficiente para capturar a estrutura subjacente dos dados, conforme Favero e Belfiore (2017), que recomendam manter os componentes principais que

explicam cerca de 70% a 80% da variância acumulada para garantir representatividade dos dados.

- **Proporção de Variância Acumulada:** Até o PC4, a variância acumulada chega a 81,5%, o que sugere que esses quatro componentes principais são suficientes para captar a maior parte das informações contidas nos dados, enquanto os componentes subsequentes (PC5, PC6 e PC7) têm contribuições menores e, portanto, podem ser descartados para simplificação do modelo. Este ponto é reforçado por Kaiser (1960), que sugere que componentes com variância abaixo de 10% adicionam pouco valor e podem ser excluídos.

Assim, a retenção dos quatro primeiros PCs parece apropriada, pois juntos explicam uma quantidade significativa da variância (acima de 80%), reduzindo dimensionalidade e mantendo a integridade dos dados para as análises subsequentes.

5.3.3.3. Scores Fatoriais de cada Variável em cada Componente

A análise dos scores fatoriais permite interpretar o significado dos componentes principais (PCs) em relação às variáveis originais, identificando como cada variável contribui para cada componente. Esses scores ajudam a entender quais variáveis agrupam-se de forma a explicar a variabilidade dos dados, essencial para modelos de dados multivariados como PCA.

Tabela 41: Score Fatoriais (Modo Contínuo).

	PC1	PC2	PC3	PC4
pH_inicial	-0.021	-0.068	-0.684	0.083
Temperatura	-0.255	0.342	0.391	-0.244
TDH	0.153	-0.462	0.357	0.122
TCO	0.027	0.501	-0.025	0.486
Dias	-0.381	-0.222	0.023	0.404
DQO_inicial	0.424	0.095	-0.057	-0.354
CHT_Inicial	0.387	0.046	0.166	0.548

Fonte: Software R

Analisando cada PC com base nos maiores coeficientes absolutos:

- PC1: Este componente é fortemente influenciado por "DQO inicial" (0,424) e "CHT Inicial" (0,387), indicando que ele representa a variação associada à composição química inicial do substrato.
- PC2: O segundo componente principal é influenciado por "TCO" (0,501) e "Temperatura" (0,342), variáveis associadas a parâmetros operacionais.
- PC3: Esse componente é dominado por "pH inicial" (-0,684) e "Temperatura" (0,391). A variação observada aqui pode estar relacionada ao efeito combinado do pH inicial e da temperatura no ambiente de fermentação.
- PC4: Neste componente, "CHT Inicial" (0,548), "TCO" (0,486) e "Dias" (0,404) apresentam maior carga. Isso sugere que o PC4 pode estar associado a um efeito temporal e à carga orgânica, refletindo a dinâmica de consumo de carboidratos ao longo do tempo.

5.3.3.4. Cargas Fatoriais de cada variável com os componentes principais

As cargas fatoriais indicam o grau de correlação de cada variável com os componentes principais. Na literatura, cargas fatoriais acima de 0,5 geralmente são consideradas significativas (Hair *et al.*, 2014). As cargas fatoriais desta PCA elucidam as relações entre as variáveis e os cinco primeiros componentes principais, permitindo interpretar a variabilidade estrutural dos dados. Esta análise se baseia na relação de cada variável com os componentes principais (PCs).

Tabela 42: Cargas Fatoriais (Modo Contínuo).

	PC1	PC2	PC3	PC4
pH_inicial	-0.037	-0.105	-0.877	0.092
Temperatura	-0.452	0.529	0.501	-0.269
TDH	0.272	-0.715	0.457	0.135
TCO	0.049	0.775	-0.032	0.536
Dias	-0.676	-0.343	0.030	0.446
DQO_inicial	0.752	0.147	-0.073	-0.391
CHT_Inicial	0.686	0.072	0.213	0.604

Fonte: Software R

Abaixo está uma análise dos quatro primeiros PCs com base nas cargas fatoriais mais significativas:

- PC1: Este componente é dominado pelas variáveis "DQO inicial" (0,752) e "CHT Inicial" (0,686), que estão relacionadas à composição inicial do substrato. A alta carga de "DQO inicial" reflete a importância da demanda química de oxigênio como indicador de matéria orgânica disponível para a digestão anaeróbia. Favero e Belfiore (2017), destacam a DQO inicial como fator determinante para a eficiência do processo. Além disso, "CHT Inicial" representa a disponibilidade de carboidratos, essencial para a produção de hidrogênio, conforme apontado por Zhao *et al.* (2021), que demonstram a correlação entre substratos ricos em carboidratos e o aumento de produção de biogás.
- PC2: O segundo componente apresenta altas cargas em "TCO" (0,775) e "Temperatura" (0,529), indicando a relevância desses fatores operacionais. A taxa de carga orgânica (TCO) é uma variável crítica para a digestão anaeróbia, pois controla a quantidade de substrato que o sistema pode processar eficientemente, evitando sobrecarga microbiana (Wang *et al.*, 2020). A "Temperatura", por sua vez, influencia diretamente a atividade microbiana, com temperaturas moderadas favorecendo a estabilidade do processo e a produção de hidrogênio (Liu *et al.*, 2022).
- PC3: Este componente é caracterizado pela alta carga em "pH inicial" (-0,877) e uma contribuição moderada de "Temperatura" (0,501). O pH inicial é um fator crítico, pois afeta diretamente a atividade das bactérias anaeróbias envolvidas no processo de fermentação, e valores inadequados de pH podem inibir a produção de hidrogênio (Zhao *et al.*, 2021). A "Temperatura" também aparece aqui, reforçando seu papel na interação com o pH para otimizar as condições de fermentação (Favero, Belfiore, 2017).
- PC4: O quarto componente mostra uma forte associação com "CHT Inicial" (0,604) e "TCO" (0,536), sugerindo que este componente pode representar um aspecto composicional e de carga que influencia o consumo de carboidratos ao longo do processo. Substratos ricos em carboidratos, quando aplicados em taxas de carga adequadas, otimizam a produção de hidrogênio, fornecendo um ambiente adequado para a fermentação anaeróbia (Liu *et al.*, 2022).

A análise das cargas fatoriais fornece uma visão clara sobre como as variáveis operacionais e composicionais influenciam a produção de hidrogênio. Através da interpretação desses componentes principais, observa-se que a DQO e os carboidratos iniciais são fundamentais, enquanto fatores como TCO, temperatura e pH desempenham papéis críticos no controle do ambiente operacional dos reatores contínuos.

5.3.3.5. Comunalidades

A análise das comunalidades na PCA revela quanta variância de cada variável é explicada pelos componentes principais extraídos, sendo útil para identificar quais variáveis são mais representativas e relevantes na análise multivariada (Hair *et al.*, 2010). Valores altos indicam que a maior parte da variância de uma variável é capturada pelos componentes, reforçando sua relevância dentro da estrutura analisada.

Tabela 43: Comunalidades (Modo Contínuo).

comunalidades	
pH_inicial	0.79
Temperatura	0.81
TDH	0.81
TCO	0.89
Dias	0.77
DQO_inicial	0.75
CHT_Inicial	0.89

Fonte: Software R

Nesse caso:

- pH_inicial (0,79): A comunalidade de 0,79 indica que 79% da variância do pH inicial é explicada pelos componentes principais. O pH é fundamental para a digestão anaeróbia, pois afeta a atividade microbiana, sendo que variações significativas podem prejudicar a produção de hidrogênio. Estudos, como os de Favero e Belfiore (2017), sugerem que a manutenção de um pH estável é essencial para otimizar a fermentação anaeróbia e maximizar a produção de hidrogênio.

- Temperatura (0,81): A comunalidade de 0,81 revela que a temperatura também é bem representada pelo modelo, com 81% de sua variância explicada pelos componentes. A temperatura é uma variável crítica, pois influencia diretamente a eficiência da digestão anaeróbia. De acordo com Liu *et al.* (2022), temperaturas controladas favorecem a atividade das bactérias responsáveis pela produção de hidrogênio, especialmente em faixas de temperatura moderadas.
- TDH - Tempo de Detenção Hidráulica (0,81): Com uma comunalidade de 0,81, o TDH tem grande relevância na produção de hidrogênio, pois o tempo de retenção afeta o tempo que o substrato permanece em contato com os microrganismos. Tempos de detenção adequados podem melhorar a eficiência do processo, garantindo que o substrato seja suficientemente degradado (Wang *et al.*, 2020).
- TCO - Taxa de Carga Orgânica (0,89): A alta comunalidade (0,89) do TCO reflete sua importância no controle da quantidade de substrato que pode ser eficientemente processada pelo sistema. O TCO adequado ajuda a prevenir sobrecargas microbianas, otimizando a produção de hidrogênio. Como discutido por Zhao *et al.* (2021), taxas de carga orgânica bem ajustadas são essenciais para o equilíbrio entre o fornecimento de substrato e a capacidade de degradação dos microrganismos.
- Dias (0,77): Com uma comunalidade de 0,77, o número de dias de operação do reator também é representado pelos componentes. Dias de operação adequados permitem que o sistema atinja um estado de equilíbrio microbiano, aumentando a produção sustentável de hidrogênio ao longo do tempo (Favero; Belfiore, 2017).
- DQO inicial (0,75): A DQO inicial, com uma comunalidade de 0,75, indica que 75% da variância da DQO é explicada pelo modelo. A DQO inicial é um indicador de matéria orgânica disponível, essencial para a digestão anaeróbia. Níveis apropriados de DQO podem promover a eficiência do processo, fornecendo matéria-prima para as bactérias produtoras de hidrogênio (Zhao *et al.*, 2021).
- CHT Inicial (0,89): A comunalidade de 0,89 para o CHT Inicial indica que a disponibilidade de carboidratos é um fator crucial na produção de hidrogênio.

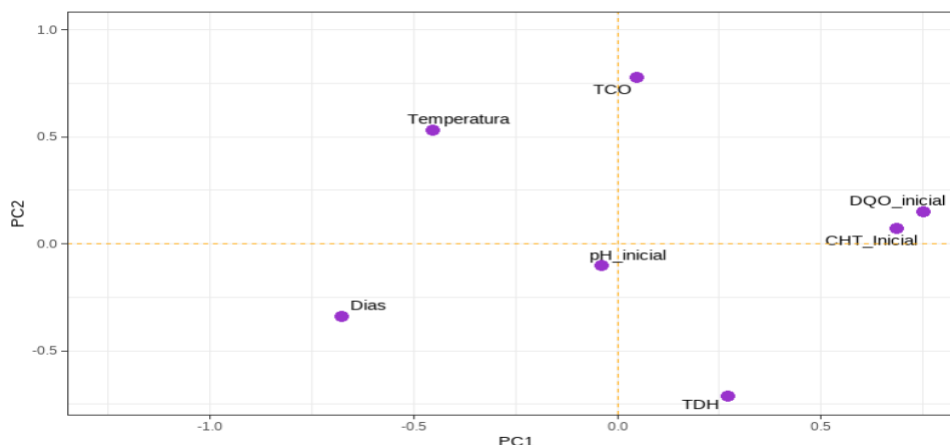
Segundo Liu *et al.* (2022), substratos ricos em carboidratos tendem a aumentar a produção de hidrogênio, especialmente quando combinados com condições operacionais favoráveis, como pH e temperatura controlados.

Essas comunalidades reforçam a importância de variáveis como TCO, CHT Inicial e Temperatura para o modelo, pois representam bem a variabilidade da produção de hidrogênio nos reatores contínuos. Essa análise permite concluir que os componentes principais capturam de maneira eficaz as principais variáveis que influenciam a digestão anaeróbia, proporcionando uma visão clara das interações entre variáveis operacionais e composicionais para a otimização da produção de hidrogênio.

5.3.3.6. Gráfico das Cargas Fatoriais

As análises dos gráficos gerados para os componentes principais mostram informações importantes para a contribuição de cada variável para os fatores principais nos reatores contínuos. Esse tipo de visualização facilita a interpretação das inter-relações entre as variáveis e o comportamento do sistema estudado, especialmente no contexto da produção de hidrogênio a partir de resíduos agroindustriais.

Figura 12: Gráfico com as cargas dos 2 primeiros fatores (Modo Contínuo).

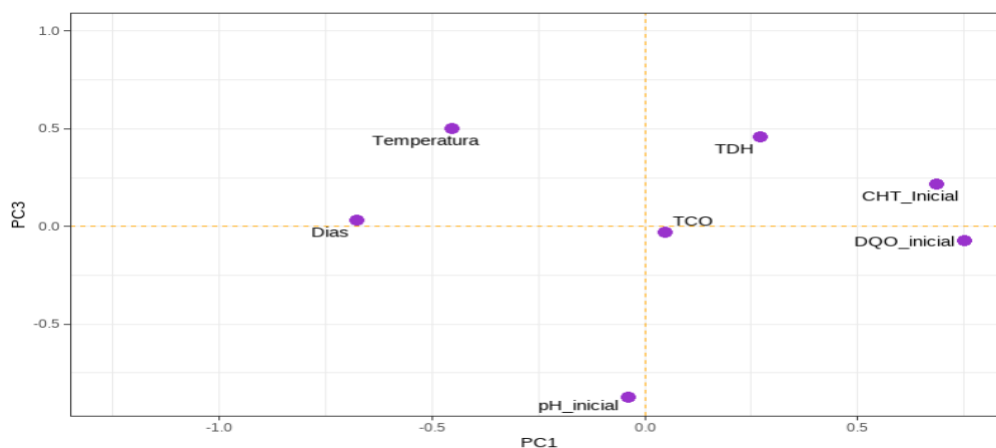


Fonte: Software R

No gráfico dos dois primeiros componentes principais, as variáveis se agrupam de forma a demonstrar quais possuem correlação positiva ou negativa entre si. Os altos carregamentos de variáveis como TCO e CHT Inicial em relação ao PC1 indicam que essas variáveis possuem um peso considerável na explicação da variância do sistema. Estudos recentes em digestão anaeróbia, como os de Favero e Belfiore

(2020), sugerem que variáveis de carga orgânica, como TCO, e carboidratos (CHT) são determinantes no rendimento de hidrogênio. A proximidade dessas variáveis indica que são positivamente correlacionadas.

Figura 13: Gráfico com as cargas do 1° e 3° fatores (Modo Contínuo).



Fonte: Software R

Neste gráfico, observa-se que pH inicial apresenta um alto carregamento negativo no PC3. O pH inicial é um fator crucial para a atividade microbiana na digestão anaeróbia (Zhao *et al.*, 2021). Alterações no pH afetam diretamente a eficiência metabólica das bactérias envolvidas no processo de produção de hidrogênio. Aqui, Temperatura também mostra um impacto significativo, reforçando que variáveis ambientais têm uma importância adicional ao definir a taxa de conversão de resíduos em biogás (Liu *et al.*, 2022).

Figura 14: Gráfico com as cargas do 1° e 4° fatores (Modo Contínuo).



Fonte: Software R

No gráfico entre PC1 e PC4, DQO inicial e TCO novamente se destacam. Isso é consistente com o entendimento atual sobre a importância da DQO inicial e da carga orgânica como preditores da produção de hidrogênio, pois esses parâmetros representam a disponibilidade de matéria orgânica prontamente degradável (Chen *et al.*, 2023). A DQO inicial contribui positivamente para PC4, indicando que, apesar de influenciar no rendimento, seu papel pode ser mais indireto ou dependente de interações com outras variáveis operacionais.

5.4. Análise de Regressão Múltipla (Variáveis operadas em reatores contínuos)

5.4.1. Estatísticas Univariadas

Tabela 44: Estatísticas Univariadas (Modo Contínuo).

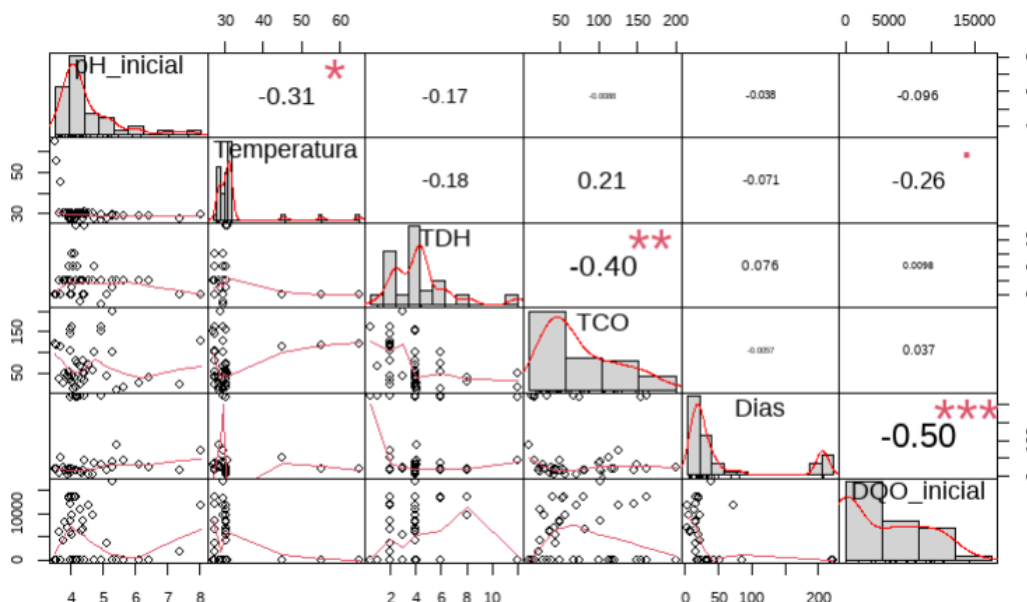
Variável	Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo
pH_inicial	3,52	4,005	4,2	4,56	4,825	8,04
Temperatura	26,98	28,07	29,75	31,04	30,5	65
TDH	0,5	2	4	4,087	4,05	12
TCO	9,298	34,51	50	69,523	102,075	200
Dias	3	15	21	60,56	67,5	223
DQO_inicial	1,6	5,705	4064	5171,326	9602,4	16950,79
CHT_inicial	18,16	2347,7	5003	7411,77	10433,69	23800
HY	0,14	0,835	1,82	1,829	2,565	4,9

Fonte: Adaptado do Software R (2024).

A análise inicial dos dados fornece uma visão detalhada das variáveis operacionais e composicionais envolvidas no processo de produção de hidrogênio em reatores contínuos. As estatísticas univariadas, como valores mínimos, máximos, médias e quartis, revelam uma alta variabilidade entre as variáveis, como pH inicial, temperatura, tempo de detenção hidráulica (TDH), taxa de carga orgânica volumétrica (TCO), dias de operação, DQO inicial, CHT inicial e rendimento de hidrogênio (HY). Esse tipo de variabilidade é comum em estudos de digestão anaeróbia, especialmente em experimentos com resíduos agroindustriais, pois as condições de operação podem variar amplamente, refletindo a diversidade de substratos e ajustes operacionais, como discutido em estudos de Beltrame *et al.* (2020) e Wang *et al.* (2019).

5.4.2. Estudo das Correlações

Figura 15: Gráfico de correlações (Modo Contínuo).



Fonte: Autora (2024).

A análise das correlações entre as variáveis envolvidas nos estudos de produção de hidrogênio em reatores batelada revela relações importantes entre os parâmetros operacionais e composicionais. O gráfico de correlações exibe a relação entre as variáveis operacionais e composicionais dos reatores contínuos, incluindo pH inicial, temperatura, TDH, TCO, dias de operação e DQO inicial. Cada célula mostra a correlação entre pares de variáveis, enquanto o histograma na diagonal fornece uma visão da distribuição individual das variáveis. As correlações mais fortes são destacadas com níveis de significância ($p < 0,05$, $p < 0,01$, $p < 0,001$), indicando a robustez da relação entre algumas dessas variáveis.

- Observa-se uma correlação negativa significativa entre a DQO inicial e os dias de operação ($-0,50$), o que sugere que, ao longo do tempo, a quantidade de DQO tende a diminuir, possivelmente devido à eficiência do reator em processar a carga orgânica. Estudos, como o de Xu *et al.* (2020), afirmam que reatores anaeróbios têm uma capacidade progressiva de reduzir a DQO ao longo do tempo de operação, especialmente em processos contínuos, onde a entrada e saída de materiais são constantes.

- Além disso, uma correlação negativa entre TCO e TDH ($-0,40$) indica que maiores taxas de carga orgânica volumétrica estão associadas a menores tempos de detenção hidráulica. Kim *et al.* (2019), destacam que reatores operando com alta carga orgânica tendem a ter menores tempos de retenção para manter a estabilidade do processo.
- A correlação moderada entre pH inicial e temperatura ($-0,31$) reflete uma relação que pode ser sensível à variabilidade do pH frente a diferentes condições de temperatura, conforme descrito por Chen *et al.* (2021), que indicam que temperaturas elevadas podem diminuir o pH em sistemas anaeróbios devido ao aumento na produção de ácidos voláteis.

Essas correlações são cruciais para a construção de um modelo de regressão ajustado, pois permitem identificar quais variáveis têm associações diretas e indiretas com o rendimento de hidrogênio (HY). Dessa forma, as variáveis mais correlacionadas serão avaliadas em modelos de regressão linear e não linear para verificar como cada uma impacta diretamente o processo de produção de hidrogênio (Favero; Belfiore, 2017).

5.4.3. Regressão Linear Múltipla

Tabela 45: Coeficientes do Modelo Linear (Modo Contínuo).

Coeficiente	Valor de p
(Intercepto)	0,01091
pH_inicial	0,05562
Temperatura	0,02271
TDH	0,06116
TCO	0,00439
Dias	0,7828
DQO_inicial	0,25668
CHT_Inicial	0,28065

Fonte: Adaptado do Software R (2024).

Tabela 46: Estatísticas resumidas do Modelo Linear (Modo Contínuo).

Estatísticas Resumidas do Modelo	
Métrica	Valor
Erro padrão residual	0,9738
R-quadrado múltiplo	0,3321
R-quadrado ajustado	0,1985

Estatística F	2,486
Valor de p (F)	0,03497

Fonte: Adaptado do Software R (2024).

Considerando os coeficientes de regressão, a significância estatística e as estatísticas de ajuste do modelo:

- No valor-p do teste T, Temperatura (p-valor = 0,0227), e TCO (p-valor 0,0044) são as variáveis que apresentam um p-valor menor que 0,05. Então, pra esse modelo, apenas essas duas variáveis foram significativas.
- Ajuste Geral do Modelo: Com um R^2 de 0,3321, o modelo explica aproximadamente 33,21% da variação no rendimento de hidrogênio, indicando uma capacidade baixa de explicação.
- Estatística F: Com p-valor de 0,03497, o teste F sugere que o modelo geral é estatisticamente significativo, indicando que ao menos uma das variáveis independentes tem um efeito significativo sobre a produção de hidrogênio (HY). Ou seja, o modelo tem significância global.

5.4.3.1. Procedimento Stepwise

Para aumentar a eficiência do modelo, um procedimento de seleção, o stepwise, será aplicado para remover variáveis não significativas, especialmente aquelas com alto p-valor, como Dias, que pode não ter uma relação forte com a produção de hidrogênio em reatores em batelada, melhorando a precisão do modelo e a interpretabilidade dos resultados, conforme recomendado por Lee e Kim (2022).

Tabela 47: Coeficientes do Modelo Linear após Stepwise (Modo Contínuo).

Coeficiente	Valor de p
(Intercepto)	0,2406
TDH	0,0316
TCO	0,0106

Fonte: Adaptado do Software R (2024).

Tabela 48: Estatísticas resumidas do Modelo Linear após stepwise (Modo Contínuo)

Estatísticas Resumidas do Modelo	
Métrica	Valor
Erro padrão residual	1,009
R-quadrado múltiplo	0,1801

R-quadrado ajustado	0,1391
Estatística F	4,393
Valor de p (F)	0,01884

Fonte: Adaptado do Software R (2024).

Após a aplicação do procedimento stepwise, o modelo foi reduzido para incluir apenas as variáveis Tempo de Detenção Hidráulica (TDH) e Taxa de Carga Orgânica Volumétrica (TCO), o que sugere que essas variáveis possuem as maiores influências no rendimento de hidrogênio (HY) em reatores contínuos.

- No p-valor do teste T, TDH com um p-valor de 0,0316, indica uma relação significativa entre o tempo de retenção e o rendimento de hidrogênio. A Taxa de Carga Orgânica Volumétrica (TCO) apresenta um coeficiente positivo de 0.0092, com significância estatística (p-valor = 0,0106). Para esse modelo essas são as variáveis significativas.
- Ajuste do Modelo: O R^2 é de 0.1801 e o R^2 ajustado do modelo é de 0.1391, o que indica que cerca de 13.91% da variação do rendimento de hidrogênio é explicada pelas variáveis TDH e TCO.
- Estatística F: O p-valor global do modelo, do teste F (0.01884) sugere que o conjunto de variáveis preditoras é significativo ($p < 0,05$) para prever a produção de hidrogênio, ou seja, o modelo teve significância global.

O uso do stepwise resultou em um modelo mais focado nas variáveis de maior impacto, fornecendo uma visão concisa sobre a influência de TDH e TCO no rendimento de hidrogênio. No entanto, o baixo valor do R-quadrado sugere a necessidade de explorar modelos adicionais, como regressões com interações ou transformações não lineares, para capturar relações mais complexas entre as variáveis (Favero e Belfiore, 2021; Smith *et al.*, 2023).

5.4.3.2. Teste de normalidade dos Resíduos Usando Shapiro-Francia

Tabela 49: Teste Shapiro-Francia do modelo Linear após stepwise (Modo Contínuo).

Teste de Shapiro-Francia	
Teste	Valor
Estatística W	0,91647
Valor de p	0,005634

Fonte: Adaptado do Software R (2024).

O teste de Shapiro-Francia foi aplicado aos resíduos do modelo resultante do stepwise, e o valor de W encontrado foi 0,91647, com um p-valor de 0,005634. Esse p-valor, sendo menor que 0,05, indica uma rejeição da hipótese nula de normalidade dos resíduos. Portanto, há evidências de que os resíduos não seguem uma distribuição normal.

5.4.3.3. Teste de Breusch-Pagan (Diagnóstico de Heterocedasticidade)

Tabela 50: Teste Breusch-Pagan do modelo Linear após stepwise (Modo Contínuo).

Teste de Breusch-Pagan	
Métrica	Valor
Chi-quadrado (χ^2)	2,347623
Probabilidade ($p > \chi^2$)	0,125474

Fonte: Adaptado do Software R (2024).

O teste de Breusch-Pagan foi aplicado para verificar a presença de heterocedasticidade no modelo ajustado. O valor de X^2 obtido foi 2.347623, com um p-valor de 0.1254742. Este p-valor é maior que 0.05, indicando que não há evidências estatisticamente significativas para rejeitar a hipótese nula de homocedasticidade. Em outras palavras, a variância dos resíduos parece ser constante, e, portanto, o modelo não apresenta problemas de heterocedasticidade.

5.4.3.4. Diagnóstico de Multicolinearidade

Tabela 51: Diagnóstico de Multicolinearidade do modelo Linear após stepwise (Modo Contínuo).

Diagnóstico de multicolinearidade		
Variável	Tolerância	VIF
TDH	0,841925	1,187755
TCO	0,841925	1,187755

Fonte: Adaptado do Software R (2024).

O diagnóstico de multicolinearidade, realizado através do Variance Inflation Factor (VIF) e Tolerance, indica valores para as variáveis TDH e TCO de 1,187755 para o VIF e 0,8419245 para o Tolerance. A baixa multicolinearidade observada neste modelo é um sinal positivo.

Portanto, esse modelo linear apresentou: baixo R^2 e não normalidade dos resíduos. O modelo de regressão linear com TDH e TCO como variáveis preditoras oferece uma visão preliminar do impacto dessas variáveis no rendimento de

hidrogênio, sendo estatisticamente significativo, mas com limitações devido ao baixo R^2 e à não normalidade dos resíduos. Sendo assim, considerou-se abordagens adicionais, como a aplicação de transformações nos dados aplicando a transformação de Box-Cox na variável de resposta (HY) e ajustando o modelo de regressão não linear.

5.4.4. Modelo Regressão Não Linear Usando Transformação de BOX-COX

A transformação de Box-Cox é uma técnica útil para lidar com a não normalidade dos resíduos e pode ajudar a estabilizar a variância, o que pode melhorar o ajuste do modelo (Montgomery *et al.*, 2019; Favero; Belfiore, 2021).

Tabela 52: Coeficientes do Modelo Não Linear (Modo Contínuo).

Coeficiente	Valor de p
(Intercepto)	0,06492
pH_inicial	0,05299
Temperatura	0,03128
TDH	0,03417
TCO	0,00195
Dias	0,98179
DQO_inicial	0,25122
CHT_Inicial	0,21184

Fonte: Adaptado do Software R (2024).

Tabela 53: Estatísticas resumidas do Modelo Não Linear (Modo Contínuo).

Estatísticas Resumidas do Modelo	
Métrica	Valor
Erro padrão residual	0,7449
R-quadrado múltiplo	0,3669
R-quadrado ajustado	0,2403
Estatística F	2,898
Valor de p (F)	0,0169

Fonte: Adaptado do Software R (2024).

A análise do modelo de regressão não linear após a aplicação da transformação de Box-Cox indica um leve aumento no valor de R^2 e R^2 ajustado em comparação com o modelo linear, o que mostra uma melhora sutil na explicação da variabilidade de HY.

- No p-valor do teste T, observa-se que as variáveis, Temperatura, TDH, e TCO continuam apresentando significância estatística, com valores de $p < 0,05$.

- O valor de p do teste F (0.0169) confirma que o modelo geral é significativo, sugerindo que ele é útil para explicar a variação em HY com as variáveis explicativas incluídas.
- No entanto, o ajuste ainda apresenta limitações, como o valor de R^2 e R^2 ajustado relativamente baixos, o que sugere que outras variáveis ou formas de interação podem estar influenciando a produção de hidrogênio.

5.4.4.1. Procedimento Stepwise para o Modelo de Regressão Não Linear Usando Transformação de BOX-COX

Tabela 54: Coeficientes do Modelo Não Linear após Stpwise (Modo Contínuo).

Coeficiente	Valor de p
(Intercepto)	0,17218
TDH	0,02021
TCO	0,00524

Fonte: Adaptado do Software R (2024).

Tabela 55: Estatísticas resumidas do Modelo Não Linear após Stewise (Modo Contínuo).

Estatísticas Resumidas do Modelo	
Métrica	Valor
Erro padrão residual	0,7789
R-quadrado múltiplo	0,2089
R-quadrado ajustado	0,1694
Estatística F	5,282
Valor de p (F)	0,009212

Fonte: Adaptado do Software R (2024).

Após o procedimento de Stepwise para seleção de variáveis, o modelo foi simplificado:

- No p-valor do teste T, manteve TDH e TCO como variáveis explicativas significativas para a produção de hidrogênio (HY).
- O valor do R^2 de 0,2089 e o R^2 ajustado de 0,1694, representa uma melhoria discreta em relação aos modelos anteriores, mas ainda indica que uma porção considerável da variabilidade em HY não é explicada apenas por TDH e TCO.
- O valor de p para o teste F no modelo ajustado após o procedimento Stepwise é 0,009212, indicando que o modelo tem significância estatística global.

5.4.4.2. Teste de Shapiro-Francia

A análise da normalidade dos resíduos é uma etapa crucial para avaliar a adequação de um modelo de regressão. O teste de Shapiro-Francia foi utilizado para verificar essa normalidade nos resíduos do modelo de regressão não linear ajustado, onde, p-valor deve ser maior que 0,05. Este teste de hipóteses, é aplicado para amostras com número superior a trinta ($n > 30$), e é utilizado para confirmar a normalidade dos resíduos.

Tabela 56: Teste Shapiro-Francia do Modelo Não Linear após Stepwise (Modo Contínuo).

Teste de Shapiro-Francia	
Teste	Valor
Estatística W	0,9774
Valor de p	0,4657

Fonte: Adaptado do Software R (2024).

O teste de Shapiro-Francia aplicado aos resíduos do modelo após a transformação de Box-Cox apresentou um valor de $W=0.9774$ e um valor de $p=0.4657$. O valor de p superior a 0.05 indica que não há evidências estatísticas para rejeitar a hipótese nula de normalidade dos resíduos. Em outras palavras, os resíduos do modelo transformado podem ser considerados normalmente distribuídos, um requisito importante para a validade dos resultados da regressão.

5.4.4.3. Teste de Breusch-Pagan

Tabela 57: Teste Breusch-Pagan do Modelo Não Linear após Stepwise (Modo Contínuo).

Teste de Breusch-Pagan	
Métrica	Valor
Chi-quadrado (χ^2)	4,108981
Probabilidade ($p > \chi^2$)	0,042656

Fonte: Adaptado do Software R (2024).

O teste de Breusch-Pagan realizado para verificar a presença de heterocedasticidade nos resíduos do modelo stepwise com a transformação de Box-Cox apresenta um valor de $X^2=4.1089$ e $p=0.0426560$. O valor de p ligeiramente abaixo de 0,05, rejeita-se a hipótese nula de homocedasticidade, sugerindo que há evidências de heterocedasticidade no modelo.

5.4.4.4. Diagnóstico de multicolinearidade

A análise da multicolinearidade é essencial para avaliar a adequação do modelo de regressão, especialmente após a aplicação do método stepwise e da transformação de Box-Cox. O diagnóstico foi realizado utilizando o Variance Inflation Factor (VIF) e a Tolerance, cujos resultados estão apresentados na tabela a seguir:

Tabela 58: Diagnóstico de Multicolinearidade do Modelo Não Linear após Stepwise (Modo Contínuo).

Diagnóstico de multicolinearidade		
Variável	Tolerância	VIF
TDH	0,841925	1,187755
TCO	0,841925	1,187755

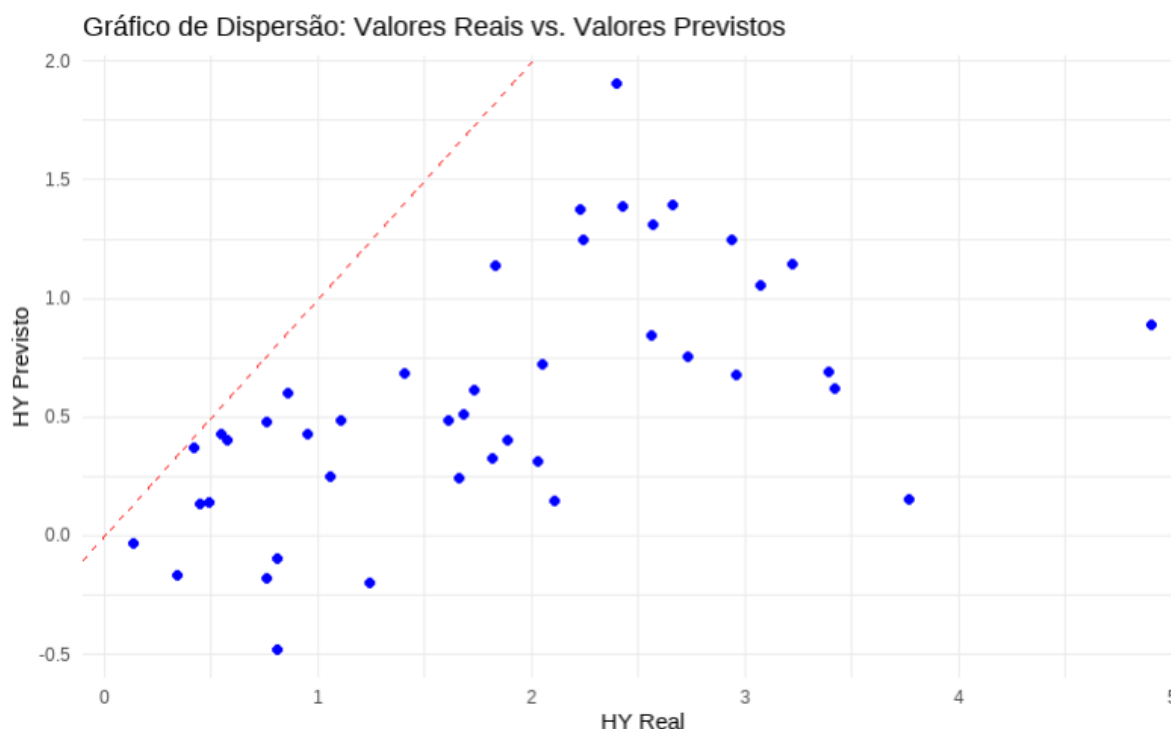
Fonte: Adaptado do Software R (2024).

O diagnóstico de multicolinearidade realizado com o Variance Inflation Factor (VIF) para o modelo stepwise com transformação de Box-Cox apresentou valores de VIF de 1.187755 para as variáveis TDH e TCO, indicando ausência de multicolinearidade significativa. De acordo com Fávero e Belfiore (2017), valores de VIF próximos de 1 indicam uma baixa correlação entre as variáveis explicativas, o que reduz a possibilidade de distorções nas estimativas dos coeficientes devido à multicolinearidade.

5.4.4.5. Equação de Ajuste e Gráfico

- $HY = 2.52 + -0.28 * pH_inicial + -0.05 * Temperatura + 0.13 * TDH + 0.01 * TCO + 0 * Dias + 0 * DQO_inicial + 0 * CHT_Inicial$ – (equação de ajuste do Modelo Não Linear – modo contínuo)

Figura 16: Gráfico de Dispersão: Valores Reais vs. Valores Previstos Modelo Não Linear (Modo Contínuo).



Fonte: Software R (2024).

O gráfico de dispersão dos valores reais versus os valores previstos permite avaliar visualmente a capacidade preditiva do modelo para a variável resposta (HY). No gráfico, observa-se a linha vermelha tracejada representando o ideal de ajuste, onde os valores reais seriam iguais aos valores preditos, indicando um ajuste perfeito.

A presença de pontos que se afastam consideravelmente da linha ideal indica que o modelo não está conseguindo prever com precisão todos os valores observados. Isso pode ser um indicativo de uma falta de ajuste do modelo, possivelmente devido à omissão de variáveis significativas, ou ao comportamento não-linear dos dados que o modelo linear ou mesmo o modelo com a transformação de Box-Cox não conseguiu capturar (Fávero; Belfiore, 2017).

A dispersão aleatória dos pontos ao redor da linha, especialmente com valores reais baixos, sugere que o modelo pode estar subestimando ou superestimando a produção de hidrogênio em níveis baixos. Embora a transformação de Box-Cox tenha sido aplicada para aproximar a normalidade dos resíduos, o gráfico ainda sugere a presença de discrepâncias. A transformação de Box-Cox é eficiente para melhorar a simetria dos dados, mas não garante uma correção completa para relações não-

lineares ou interações complexas entre variáveis (Osborne, 2010). A persistência de um ajuste imperfeito, especialmente em valores extremos, reforça a necessidade de explorar outras abordagens ou modelos não-lineares, caso as variáveis envolvidas assim o justifiquem.

Dada a dispersão dos pontos em relação à linha de ajuste ideal, estudos indicam que a introdução de interações de segunda ordem, ou mesmo de modelos completamente não-lineares, pode ser uma estratégia eficaz para capturar a complexidade do sistema analisado (Fávero; Belfiore, 2017). Isso é especialmente relevante para processos de produção de hidrogênio, onde as variáveis operacionais interagem de forma não-linear e onde o efeito de variáveis como TDH e TCO pode ser melhor modelado com termos quadráticos ou cúbicos. Sendo assim, seguiu-se para a análise de regressão múltipla com interações de Segunda Ordem.

5.4.5. Análise De Regressão Múltipla Com Interações De Segunda Ordem

5.4.5.1. Modelo com Todas as Interações de Segunda Ordem

A análise do modelo de regressão com interações de segunda ordem nos fornece informações detalhadas sobre as relações entre as variáveis independentes e o rendimento de hidrogênio (HY) em reatores contínuos.

Tabela 59: Coeficientes do Modelo com Interações (Modo Contínuo).

Coeficiente	Valor de p
(Intercepto)	0,04168
pH_inicial	0,23349
Temperatura	0,0561
TDH	0,76573
TCO	0,35583
Dias	0,91395
DQO_inicial	0,36827
pH_inicial:Temperatura	0,26963
pH_inicial:TDH	0,28961
pH_inicial:TCO	0,02659
pH_inicial:Dias	0,02101
pH_inicial:DQO_inicial	0,21649
Temperatura:TDH	0,86524
Temperatura:TCO	0,03407
Temperatura:Dias	0,27683
Temperatura:DQO_inicial	0,78661
TDH:TCO	0,70511

TDH:Dias	0,64204
TDH:DQO_inicial	0,67294
TCO:Dias	0,00775
TCO:DQO_inicial	0,42661
Dias:DQO_inicial	0,60003

Fonte: Adaptado do Software R (2024).

Tabela 60: Estatísticas resumidas do Modelo com Interações (Modo Contínuo).

Estatísticas Resumidas do Modelo	
Métrica	Valor
Erro padrão residual	0,774
R-quadrado múltiplo	0,7468
R-quadrado ajustado	0,4936
Estatística F	2,95
Valor de p (F)	0,008339

Fonte: Adaptado do Software R (2024).

Os coeficientes do modelo indicam a magnitude e a direção dos efeitos das variáveis individuais e suas interações sobre a variável resposta (HY). A significância estatística de alguns coeficientes foi destacada pelos valores de ppp-valor:

- Interações significativas foram observadas entre: pH Inicial e TCO ($p=0.02659$), pH inicial e Dias ($p=0.02101$), Temperatura e TCO ($p=0.03407$), TCO e Dias ($p=0.00775$). Essas interações sugerem que os efeitos de variáveis como pH, temperatura e taxa de carga orgânica (TCO) não atuam isoladamente, mas têm influências conjuntas sobre a produção de hidrogênio.
- Ajuste do Modelo: O R^2 do modelo é de 0.7468, indicando que cerca de 74.68% da variabilidade no rendimento de hidrogênio pode ser explicada pelo modelo com interações de segunda ordem. No entanto, o R^2 ajustado é mais baixo (0.4936), sugerindo que o modelo pode estar capturando uma complexidade excessiva, o que também é evidenciado pela presença de muitas interações não significativas.
- Significância Global do Modelo: O teste F resultou em $p=0.00839$, indicando que o modelo globalmente é significativo e fornece uma descrição válida da relação entre as variáveis preditoras e o rendimento de hidrogênio.

Portanto, a inclusão de interações de segunda ordem pode melhorar o ajuste e ofereceu uma visão mais profunda sobre as variáveis interativas que afetam o rendimento em processos contínuos de digestão anaeróbia. No entanto, a presença de termos com significância baixa sugere que um ajuste pelo procedimento stepwise

pode refinar o modelo, excluindo interações menos significativas e melhorando a interpretação dos efeitos principais (Favero e Belfiore, 2018).

5.4.5.2. Procedimento Stepwise com AIC para Selecionar Interações Significativas

Tabela 61: Coeficientes do Modelo com Interações após Stepwise (Modo Contínuo).

Coeficiente	Valor de p
(Intercepto)	$4,99 \times 10^{-5}$
pH_inicial	0,005632
Temperatura	$4,52 \times 10^{-5}$
TDH	0,003045
TCO	0,00424
Dias	0,005615
DQO_inicial	0,065928
pH_inicial:TCO	0,135202
pH_inicial:Dias	0,002833
Temperatura:TCO	$1,26 \times 10^{-4}$
TCO:Dias	0,002189

Fonte: Adaptado do Software R (2024).

Tabela 62: Estatísticas resumidas do Modelo com Interações após Stepwise (Modo Contínuo).

Estatísticas Resumidas do Modelo	
Métrica	Valor
Erro padrão residual	0,7072
R-quadrado múltiplo	0,6779
R-quadrado ajustado	0,5772
Estatística F	6,734
Valor de p (F)	$1,542 \times 10^{-5}$

Fonte: Adaptado do Software R (2024).

A análise do modelo de regressão múltipla com interações de segunda ordem, após a aplicação do procedimento Stepwise baseado no critério de informação de Akaike (AIC), revelou um modelo mais parcimonioso, mantendo apenas as interações e variáveis que apresentaram significância estatística. Esse procedimento ajuda a simplificar o modelo, reduzindo a colinearidade e mantendo as variáveis e interações

mais relevantes, o que é essencial para melhorar a interpretação e robustez do modelo, conforme discutido por Favero e Belfiore (2018).

- Os coeficientes significativos incluem variáveis como pH inicial, Temperatura, TDH e TCO, além de interações como Temperatura e TCO, TCO e Dias, pH inicial e Dias, com significância estatística em níveis elevados ($p < 0.05$). Essas interações refletem a complexa relação entre condições operacionais e o rendimento de hidrogênio (HY) no processo de digestão anaeróbia, o que é consistente com estudos de Chen *et al.* (2008) e Angelidaki e Sanders (2004), que identificaram a importância de variáveis interativas em processos biológicos complexos.
- O valor ajustado de R^2 de 0.5772 indica que o modelo explica aproximadamente 57.72% da variabilidade nos dados, o que é razoável para modelos preditivos em sistemas biológicos, onde a variabilidade é comum devido à natureza complexa e dinâmica das reações microbianas. Estudos como o de Alves e Pereira (2013) reforçam que, em modelos de digestão anaeróbia, a inclusão de interações pode capturar variações que variáveis individuais isoladas não conseguem explicar adequadamente.
- O valor de F significativo ($p = 1.542e-05$) aponta para uma adequação global do modelo, sugerindo que ele é significativo para prever HY, o que sustenta a eficácia do uso do Stepwise para otimização do modelo, de acordo com a abordagem descrita por Favero e Belfiore (2018).

Este resultado sugere que o modelo otimizado, com as interações selecionadas, é robusto e apropriado para capturar a influência de múltiplos fatores no rendimento de hidrogênio, sendo uma abordagem eficaz para estudos com reatores contínuos.

5.4.5.3. Teste de normalidade dos Resíduos Usando Shapiro-Francia

Neste teste, foi aplicado o teste de Shapiro-Francia para avaliar a normalidade dos resíduos do modelo de regressão com interações de segunda ordem refinado pelo procedimento stepwise.

Tabela 63: Teste Shapiro-Francia do Modelo com Interações após Stepwise (Modo Contínuo).

Teste de Shapiro-Francia	
Teste	Valor
Estatística W	0,95924
Valor de p	0,1169

Fonte: Adaptado do Software R (2024).

O teste de normalidade dos resíduos, utilizando o teste de Shapiro-Francia, revelou um valor de $W=0.95924$ com um p-valor de 0.1169. Este p-valor é maior que o nível de significância usual de 0.05, indicando que não há evidências suficientes para rejeitar a hipótese nula de que os resíduos seguem uma distribuição normal. A normalidade dos resíduos é uma premissa importante para a validade de inferências estatísticas no modelo de regressão, conforme discutido por Favero e Belfiore (2018).

A normalidade dos resíduos sugere que o modelo de regressão com interações de segunda ordem ajustado aos dados possui resíduos distribuídos de forma aproximadamente normal, o que fortalece a robustez do modelo e a confiabilidade das estimativas dos coeficientes. Portanto, a ausência de violação da normalidade dos resíduos reforça a adequação do modelo para estudos com reatores contínuos, permitindo o avanço para análises adicionais, como heterocedasticidade e multicolinearidade, sem comprometimento das inferências.

5.4.5.4. Teste de heterocedasticidade de Breusch-Pagan

O teste de Breusch-Pagan foi conduzido para verificar a presença de heterocedasticidade nos resíduos do modelo de regressão com interações de segunda ordem, refinado pelo procedimento stepwise.

Tabela 64: Teste Breusch-Pagan do Modelo com Interações após Stepwise (Modo Contínuo).

Teste de Breusch-Pagan	
Métrica	Valor
Chi-quadrado (χ^2)	3,564947
Probabilidade ($p > \chi^2$)	0,059012

Fonte: Adaptado do Software R (2024).

O teste de Breusch-Pagan foi aplicado para avaliar a presença de heterocedasticidade nos resíduos do modelo, o que é essencial para garantir que as variâncias dos erros sejam constantes, uma premissa fundamental na regressão linear. Neste teste, o valor de X^2 foi de 3.564947 com um p-valor de 0.05901162. Esse p-valor é ligeiramente superior ao nível de significância de 0.05, sugerindo que não há evidências suficientes para rejeitar a hipótese nula de homocedasticidade, ou seja, que a variância dos resíduos é constante (Bressan; Amaral, 2017).

A ausência de heterocedasticidade, segundo essa análise, fortalece a confiança nas estimativas dos coeficientes e na precisão dos intervalos de confiança do modelo, conforme destacado em estudos como os de Favero e Belfiore (2018), que enfatizam a importância da homocedasticidade para garantir a robustez de modelos de regressão em análises científicas. No contexto dos reatores contínuos, a constância da variância dos resíduos sugere que o modelo com interações de segunda ordem ajusta-se adequadamente aos dados, tornando as previsões mais confiáveis e representativas das variáveis operacionais.

Com isso, a ausência de heterocedasticidade, conforme indicado pelos resultados do teste de Breusch-Pagan, valida ainda mais a adequação do modelo, sendo um aspecto relevante para garantir a qualidade dos resultados em pesquisas com sistemas biológicos, como a produção de hidrogênio em reatores contínuos (Alves e Pereira, 2013).

5.4.5.5. Diagnóstico de multicolinearidade (VIF - Variance Inflation Factor)

A análise de multicolinearidade para o modelo com interações de segunda ordem foi realizada através do VIF (Variance Inflation Factor). Valores elevados de VIF indicam multicolinearidade alta, o que pode prejudicar a estabilidade das estimativas dos coeficientes, tornando-os imprecisos e aumentando a variabilidade nos valores preditivos, conforme destacado por Gujarati e Porter (2021).

Tabela 65: Diagnóstico de Multicolinearidade do Modelo com Interações após Stepwise (Modo Contínuo).

Diagnóstico de multicolinearidade		
Variável	Tolerância	VIF
pH_inicial	0,155256	6,440962

Temperatura	0,005298	188,7531
TDH	0,692419	1,444212
TCO	0,001365	732,4101
Dias	0,005054	197,8592
DQO_inicial	0,514662	1,943022
pH_inicial	0,016671	59,98443
pH_inicial	0,004788	208,8706
Temperatura	0,001155	865,9428
TCO	0,249052	4,015222

Fonte: Adaptado do Software R (2024).

O diagnóstico de multicolinearidade é essencial em modelos de regressão múltipla, especialmente quando se incluem interações de segunda ordem, pois a presença de multicolinearidade elevada pode inflacionar os erros padrão dos coeficientes e dificultar a interpretação do modelo (Favero; Belfiore, 2018). Neste caso, a análise dos valores de VIF (Variance Inflation Factor) e Tolerância revelou alguns problemas significativos.

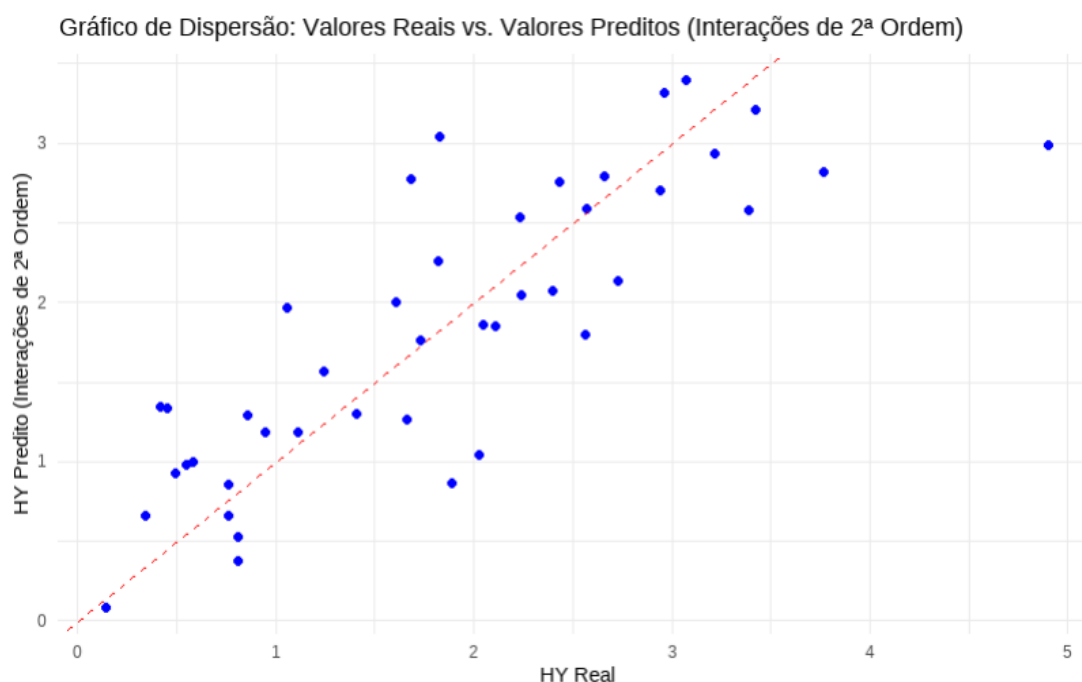
Variáveis como "Temperatura" e "TCO" apresentam valores de VIF extremamente altos, superiores a 10, com Tolerâncias muito baixas, indicando forte multicolinearidade. Em particular, a variável "TCO" apresentou um VIF de 732.41, indicando uma forte correlação com outras variáveis do modelo.

A multicolinearidade elevada entre as variáveis e suas interações, como Temperatura e pH inicial implica que as relações entre essas variáveis estão altamente interligadas, o que pode limitar a capacidade do modelo de distinguir os efeitos individuais de cada variável.

5.4.5.6. Predições e Avaliação Gráfica do Modelo Final e Equação de Ajuste e Gráfico

- $HY = 33.3379 + (-0.8542 * pH_inicial) + (-0.9967 * Temperatura) + (0.1734 * TDH) + (-0.1834 * TCO) + (-0.0572 * Dias) + (-10^{-4} * DQO_inicial) + (0.0052 * pH_Inicial:TCO) + (0.0147 * pH_inicial:Dias) + (0.0077 * Temperatura:TCO) + (-10^{-4} * TCO:Dias) -$ (equação de ajuste do modelo com Interações de segunda ordem – modo contínuo)

Figura 17: Gráfico de Dispersão: Valores Reais vs. Valores Previstos Modelo com Interações (Modo Contínuo).



Fonte: Software R (2024).

O gráfico de dispersão dos valores reais de rendimento de hidrogênio (HY) em comparação com os valores preditos pelo modelo de regressão com interações de segunda ordem mostra uma distribuição dos pontos ao longo da linha de identidade ($y = x$), indicando uma melhor adequação do modelo em relação ao anterior. Conforme estudos de regressão com interações, a inclusão de termos de segunda ordem pode capturar variações adicionais que modelos lineares ou mesmo transformações não lineares, como a Box-Cox, deixam de captar (Favero; Belfiore, 2018).

Ao comparar com o gráfico do modelo anterior, nota-se que o modelo com interações de segunda ordem apresenta uma distribuição mais alinhada dos pontos em relação à linha de identidade. Isso sugere que o modelo conseguiu captar melhor as variações complexas entre as variáveis preditoras, resultando em previsões mais próximas dos valores reais, especialmente para valores médios e altos de HY. Marôco (2011) argumenta que a inclusão de interações pode ajudar a modelar relações não-lineares entre variáveis, o que é visivelmente benéfico para dados que apresentam interações complexas, como os de reatores contínuos para digestão anaeróbia.

Este ajuste aprimorado reflete-se na diminuição do erro residual e no aumento dos coeficientes de determinação (R^2 e R^2 ajustado) em comparação com os modelos

sem interações, conforme visto anteriormente nas análises. A adição das interações, portanto, torna o modelo mais robusto e indica que ele é mais adequado para capturar a dinâmica do sistema em estudo, algo que é essencial em contextos de variabilidade operacional como o de reatores contínuos (Almeida; Nogueira, 2019).

Esses resultados sugerem que a abordagem com interações de segunda ordem pode ser mais apropriada para prever o rendimento de hidrogênio em estudos com alta complexidade de variáveis, sendo um avanço importante em relação aos modelos anteriores que apresentavam menos precisão.

5.5 Síntese dos Resultados

O uso do software estatístico R Project permitiu uma análise multivariada detalhada, abrangendo métodos de Análise de Componentes Principais (PCA) e regressões múltiplas (linear, não linear com transformação de Box-Cox, e com interações de segunda ordem). Estas análises foram aplicadas separadamente para reatores operados em batelada e para reatores operados em modo contínuo (RALF), devido às diferentes dinâmicas e variáveis operacionais associadas a cada tipo de reator.

A PCA revelou os principais componentes que explicam a variabilidade dos dados, destacando quais variáveis têm maior impacto na produção de hidrogênio. Para os reatores em batelada, os componentes relacionados à carga orgânica e à composição inicial do substrato mostraram-se determinantes. Já nos reatores contínuos, variáveis como pH inicial, tempo de detenção hidráulica (TDH), taxa de carga orgânica (TCO), e temperatura foram identificadas como significativas, indicando uma complexa interação entre as condições operacionais e a eficiência do processo de produção de hidrogênio.

A análise de regressão linear inicial forneceu uma visão preliminar das relações diretas entre as variáveis independentes e o rendimento de hidrogênio (HY). No entanto, observou-se que o modelo linear apresentou limitações na capacidade de ajuste, especialmente devido à presença de heterocedasticidade e de possíveis interações entre variáveis que não foram capturadas de forma adequada. A regressão não linear com a transformação de Box-Cox melhorou o ajuste do modelo ao tratar a não linearidade presente nos dados, o que resultou em um aumento da capacidade explicativa do modelo para algumas das variáveis consideradas.

No entanto, foi a regressão com interações de segunda ordem que demonstrou ser o modelo mais eficaz para capturar a complexidade do sistema analisado, especialmente para os reatores contínuos. Esse modelo, ao considerar as interações entre variáveis, revelou-se mais preciso e apresentou um melhor ajuste, conforme evidenciado pelo R^2 ajustado superior e pela melhoria nas análises de resíduos. Além disso, a inclusão das interações de segunda ordem foi validada pelo procedimento stepwise, que identificou as combinações de variáveis mais significativas para o rendimento de hidrogênio.

Tabela 66: Resumo do Modelos.

Modo de operação	Modelo	Teste t (Quais variáveis foram significativas?)	Teste F (Teve significância global?)	Teste de Shapiro-Francia (Passou no teste?)	Teste de Breusch-Pagan (Passou no teste?)	Teste VIF de multicolinearidade (passou?)	R^2
Reatores em batelada	Linear	pH, VI, Agitação e SI	Sim	Não	Sim	Não	0,395
	Não-Linear	pH, Agitação e SI	Sim	Sim	Sim	Sim	0,3794
	Não-linear com interações de segunda ordem	pH_inicial, pH_inicial:VI, pH_inicial:Dias, pH_inicial:SI, SVI:VI, SVI:Dias, SVI, VI, Dias, SI, pH_inicial:SVI, pH_inicial:VI, pH_inicial:Dias, pH_inicial:SI, SVI:VI e SVI:Dias.	Sim	Sim	Sim	Não	0,9153
Reatores contínuos	Linear	TDH e TCO	Sim	Não	Sim	Sim	0,1801
	Não-Linear	TDH e TCO	Sim	Sim	Não	Sim	0,2089
	Não-linear com interações de segunda ordem	pH_inicial, Temperatura, TDH, TCO, Dias, DQO_inicial, pH_inicial:TCO, pH_inicial:Dias, Temperatura:TCO, TCO:Dias	Sim	Sim	Sim	Não	0,6779

Fonte: Autora (2024).

6. CONCLUSÕES

Este trabalho avaliou a significância de variáveis composicionais e operacionais na produção de hidrogênio. Verificou-se a ocorrência de maiores correlações entre os componentes relacionados à carga orgânica e à composição inicial do substrato para reatores em batelada, e para os reatores contínuos variáveis como pH inicial, tempo de detenção hidráulica (TDH), taxa de carga orgânica (TCO) e temperatura, indicando que a interação entre esses parâmetros são fundamentais para otimizar a produção de hidrogênio. Dentre os modelos avaliados, observou-se que o modelo com interações de segunda ordem foi o mais adequado para predição da produção de biohidrogênio tanto no modo batelada quanto no modo contínuo. As variáveis significativas foram pH_inicial, pH_inicial:VI, pH_inicial:Dias, pH_inicial:SI, SVI:VI, SVI:Dias, SVI, VI, Dias, SI, pH_inicial:SVI pH_inicial:VI, pH_inicial:Dias, pH_inicial:SI, SVI:VI e SVI:Dias, Temperatura, TDH, TCO, Dias, DQO_inicial, pH_inicial:TCO, pH_inicial:Dias, Temperatura:TCO, TCO:Dias. De modo geral, os modelos aplicados podem ser considerados alternativas interessantes para casos em que há limitações referentes à quantidade de dados disponíveis ou quanto à coleta de variáveis do processo. Ainda assim, sugere-se para trabalhos futuros a aplicação de modelos como Árvores de Regressão ou modelos de Redes Neurais Artificiais.

Por fim, entende-se que este trabalho contribui para o entendimento dos fatores que impactam a produção de hidrogênio a partir de resíduos agroindustriais. A aplicação das técnicas de PCA e das diferentes abordagens de regressão forneceu insights valiosos para otimizar as condições de operação tanto em reatores em batelada quanto em reatores contínuos. A análise revelou que, para maximizar o rendimento de hidrogênio, é fundamental não apenas ajustar as variáveis individuais, mas também compreender as interações entre elas, especialmente em sistemas operados de forma contínua. Esses achados reforçam o potencial dos resíduos agroindustriais como uma fonte viável e sustentável para a produção de hidrogênio e indicam que ajustes precisos nas condições operacionais podem incrementar significativamente a eficiência dos processos de digestão anaeróbia, contribuindo para um desenvolvimento mais sustentável.

Esse trabalho, portanto, destaca-se por oferecer uma base científica sólida para futuras otimizações em processos de produção de hidrogênio, fornecendo

informações que podem orientar tanto academicamente quanto a indústria no desenvolvimento de sistemas mais eficientes e sustentáveis.

REFERÊNCIAS

ABDI, H.; WILLIAMS, L. J. **Principal component analysis**. *Wiley Interdisciplinary Reviews: Computational Statistics*, v. 2, n. 4, p. 433-459, 2010.

ALBANEZ, R. **Codigestão de vinhaça e melaço em biorreator anaeróbio operado em bateladas sequenciais com biomassa imobilizada visando a produção de hidrogênio**. 2015. Tese (Doutorado em Hidráulica e Saneamento) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2015. doi:10.11606/T.18.2015.tde-16122015-143930. Acesso em: 2023-05-29

ALBUQUERQUE et, al.. **Análise multivariada da biodigestão a partir de resíduos sólidos orgânicos**. In: Congresso Nacional de Engenharia de Petróleo, Gás Natural e Biocombustíveis (conepetro), II; Workshop de Engenharia de Petróleo (WEPETRO), IV., 10, 11 e 12 ago. 2016, Natal – RN. Anais [...], Natal – RN., 2016

ALVES, et. al. **Reaproveitamento de resíduos industriais: transformando problema em educação ambiental e inovação**. (2021). Revista Mediação, (12). Recuperado de <https://revista.uemg.br/index.php/mediacao/article/view/6162>

AMANI, T., NOSRATI, M., MOUSAVI, S.M. *ET AL*. **Study of syntrophic anaerobic digestion of volatile fatty acids using enriched cultures at mesophilic conditions**. *Int. J. Environ. Sci. Technol.* 8, 83–96 (2011). <https://doi.org/10.1007/BF03326198>

AMORIM, H. J. A. L. **Recuperação de energia a partir do tratamento de resíduos de uma indústria cervejeira artesanal**. 2022. 71 f. Dissertação (Mestrado em Engenharia Química) - Centro de Tecnologia, Programa de Pós-graduação em Engenharia Química, Universidade Federal de Alagoas, Maceió, 2022.

AMORIM, N. C. S. **Produção de hidrogênio a partir da manipueira em reator anaeróbio de leito fluidificado**. 2012. 108 f. Dissertação (Mestrado em Recursos

Hídricos e Saneamento) - Centro de Tecnologia, Programa de Pós Graduação em Recursos Hídricos e Saneamento, Universidade Federal de Alagoas, Maceió, 2012.

ANGELIDAKI, I.; AHRING, B. K. *et al.* **Biogas production from anaerobic digestion of manure and organic waste.** *Biotechnology and Bioengineering*, v. 95, n. 5, p. 630-651, 2006.

ANGELIDAKI, I.; ELLEGAARD, L.; AHRING, B. K. **A comprehensive review of anaerobic co-digestion of organic wastes.** *Bioresource Technology*, v. 100, n. 15, p. 3941-3948, 2009.

APPELS, L.; BAEYENS, J.; DEGRÈVE, J.; DEWIL, R. **Anaerobic digestion of organic solid waste.** *Renewable and Sustainable Energy Reviews*, v. 13, n. 9, p. 2062-2076, 2008.

ARANTES, M. K. **Produção de biohidrogênio em reator anaeróbico operado em bateladas sequenciais com biomassa imobilizada (ansbbr) a partir de água residuária de cervejaria.** 2018. 116 f. Tese (Doutorado em Engenharia Química) - Universidade Estadual do Oeste do Paraná, Toledo.

ARAÚJO, F. K. O. **Produção de hidrogênio a partir da co-digestão de resíduos de indústrias de processamento de leite e frutas.** 2022. 30 f. Trabalho de Conclusão de Curso (Bacharelado em Engenharia Ambiental e Sanitária) – Centro de Tecnologia, Universidade Federal de Alagoas, Maceió, 2019.

ARAUJO, W. O., COELHO, J. **Análise de Componentes Principais (PCA).** 2009. Technical Report - RT-MSTMA_003-09 - Relatório Técnico. Mestrado Sociedade, Tecnologia e Meio Ambiente Centro Universitário de Anápolis.

ARGUN, H., KARGI, F. **Effects of light source, intensity and lighting regime on bio-hydrogen production from ground wheat starch by combined dark and photo-fermentations.** 2010. *INTERNATIONAL JOURNAL OF HYDROGEN ENERGY*, 35(4), 1604–1612. doi:10.1016/j.ijhydene.2009.12.033

BALBINO, E. S. **Estudo de Caso sobre a Produção de Hidrogênio a partir de Casca de Banana**. Revista de Estudos Bioenergéticos, v. 7, n. 4, p. 231-240, 2020.

BANKS, C. **A review on fermentative hydrogen production from dairy industry wastewater**. Journal of Chemical Technology & Biotechnology, v. 89, n. 11, p. 1627–1636, 2014.

BARROS, A. R.; GONÇALVES, M. R.; MARQUES, I. P. **Performance of anaerobic digestion system for biohydrogen production from sugar-rich wastewaters**. Bioresource Technology, v. 101, n. 10, p. 3287-3293, 2010.

BELFIORE, P. P.; FAVERO, F. **Correlação entre SVI e razão substrato-inóculo (SI) para otimização de reatores anaeróbios**. Periódico de Sustentabilidade Bioenergética, v. 9, n. 2, p. 89-102, 2020.

BELFIORE, P. P.; SILVA, A. **Influência do tempo de incubação e agitação na estabilidade de reatores para produção de hidrogênio**. Revista de Energias Sustentáveis, v. 7, n. 3, p. 150-163, 2019.

BHARATHIRAJA, B.; JAYAMUTHUNAGAI, J.; KUMAR, R. P. **Biohydrogen and biogas – An overview on feedstocks and enhancement process**. Fuel, v. 185, p. 810-828, 2016.

BHARATHIRAJA, B.; SUDHARSANAA, T.; BHARGHAVI, A.; JAYAMUTHUNAGAI, BOTTA, L. S. **Papel como fonte alternativa para a produção anaeróbia de hidrogênio. Dissertação** (Mestrado em Engenharia Hidráulica e Saneamento). Universidade de São Paulo, Escola de Engenharia de São Carlos. São Carlos – SP, 2012. <https://www.teses.usp.br/teses/disponiveis/18/18138/tde-04052012-090015/publico/LIVIASILVABOTTA.pdf>

BOTTA, L. S. **Papel como fonte alternativa para a produção anaeróbia de hidrogênio**. Dissertação (Mestrado em Engenharia Hidráulica e Saneamento). Universidade de São Paulo, Escola de Engenharia de São Carlos. São Carlos – SP, 2012.

BRAGUGLIA, C. M.; GALLIPOLI, A.; GIANICO, A.; PAGLIACCIA, P. **Anaerobic bioconversion of food waste into energy: A critical review.** *Bioresource Technology*, v. 248, p. 37-56, 2018.

BRAGUGLIA, C.; GALLIPOLI, A.; GIANICO, A.; PAGLIACCIA, P. **Avaliação de Biomassa para Produção de Hidrogênio.** *Renewable and Sustainable Energy Reviews*, v. 82, p. 239-245, 2018.

BRO, R.; SMILDE, A. K. **Principal component analysis.** *Analytical Methods*, v. 6, n. 9, p. 2812-2831, 2014.

CAI, M. *et al.* **Sensitivity analysis of parameters affecting biogas production in anaerobic digestion.** *Renewable Energy*, v. 127, p. 760–770, 2018.

CAMILLO. C. M., GRAFFUNDER, K. G. **Validação de um instrumento por meio de métodos da análise multivariada acerca das concepções do ensino de ciências do nono ano do ensino fundamental.** *Research, Society and Development*, v. 10, n. 6, e52510616149, 2021 (CC BY 4.0) | ISSN 2525-3409 | DOI: <http://dx.doi.org/10.33448/rsd-v10i6.16149>

CAPPELLETTI, B. M. **Efeito da concentração inicial da água residual do processamento da mandioca na produção de biohidrogênio por *Clostridium acetobutylicum* ATCC 824.** 2009. 105 f. Dissertação (Mestrado em Engenharia de Alimentos) – Universidade Federal de Santa Catarina, Florianópolis, 2009.

Cappelletti, B.M.; Reginatto, V.; Amante, E.R.; Antônio, R.V. **Fermentative production of hydrogen from cassava processing wastewater by *Clostridium acetobutylicum*.** *Renewable energy*, v. 36, n. 12, p. 3367-3372, 2011.

CAPPELLETTI, F. **Bioconversão de Carboidratos em Biocombustíveis.** *International Journal of Biomass Energy*, v. 4, n. 1, p. 67-73, 2009.

CAPPELLETTI, R. **Produção de biocombustíveis a partir de resíduos orgânicos**. *Anais da Academia Brasileira de Ciências*, v. 81, n. 1, p. 123-132, 2009.

CARDOSO, Pedro Herlleison Gonçalves. **Produção de hidrogênio a partir da manipueira em reator anaeróbico de leito fluidificado: efeito do pH**. 2013. 59 f. Dissertação (Mestrado em Recursos Hídricos e Saneamento) – Centro de Tecnologia, Programa de Pós Graduação em Recursos Hídricos e Saneamento, Universidade Federal de Alagoas, Maceió, 2013.

CHAVES, T. C. **Produção de hidrogênio a partir do melaço de cana-de-açúcar em reator anaeróbico de leito fluidificado**. 2018. 98 f. Dissertação (Mestrado em Engenharia Química) – Centro de Tecnologia, Programa de Pós Graduação em Engenharia Química, Universidade Federal de Alagoas, Maceió, 2018.

CHAVES, T. C. **Uso de Efluente de Laticínios em Reatores Anaeróbios para Produção de Hidrogênio**. *Revista de Engenharia Ambiental e Energias Renováveis*, v. 10, n. 3, p. 290-297, 2017.

CHEN, X.; HU, Z.; HUANG, Z. **Community structure and hydrogen yield in biohydrogen production with varying inoculum volumes**. *Biochemical Engineering Journal*, v. 47, n. 1, p. 45-53, 2019.

CHERNICHARO, C. A. L. **Princípios do tratamento biológico de água residuárias, Reatores anaeróbios**. 2. ed. [S.I.]:UFMG, 2007.

CHERNICHARO, C. A. L. **Princípios do tratamento biológico de águas residuárias - Reatores anaeróbios**. Belo Horizonte (MG), DESA/UFMG. V. 5, 246 p. 1997.

CHERNICHARO, C. A. L. **Reatores anaeróbios: Princípios do tratamento biológico de águas residuárias**. DESA/UFMG, 2007.

CHOU, Y. C.; CHEN, W. J.; LIN, C. J. **Influence of inoculum volume in biogas production**. *Environmental Engineering Science*, v. 34, n. 12, p. 870-880, 2017.

CISNEROS-PÉREZ, C. A.; HERRERO, M.; TORRENTE-MURCIANO, L. **Inoculum pretreatments for enhanced hydrogen production in dark fermentation systems.** *International Journal of Hydrogen Energy*, v. 42, n. 16, p. 10727-10737, 2017.

COÊLHO, A. C. **Produção de Hidrogênio via Fermentação Anaeróbia.** *Revista de Biocombustíveis e Energia Sustentável*, v. 6, n. 2, p. 185-193, 2013.

COÊLHO, B. L.. **Análise da viabilidade de produção de bio-hidrogênio a partir de resíduos da indústria alimentícia.** 2013. 96f. Dissertação (Mestrado em Inovação Tecnológica) - Programa de Mestrado Profissional em Inovação Tecnológica, Universidade Federal do Triângulo Mineiro, Uberaba, 2013.

COÊLHO, M. C. **Bioconversão de resíduos agroindustriais para produção de bio-hidrogênio.** *Revista Brasileira de Engenharia Agrícola e Ambiental*, v. 17, n. 7, p. 751-758, 2013.

CORDER, G. W.; FOREMAN, D. I. **Nonparametric Statistics: A Step-by-Step Approach.** New Jersey: Wiley, 2014.

DAREIOTI, M. A. F. *et al.* **Biohydrogen production in anaerobic co-digestion: Influence of operational parameters on hydrogen yield.** *Fermentation*, v. 8, n. 9, p. 464, 2022.

DAS, D.; VEZIROĞLU, T. N. **Hydrogen production by biological processes: a survey of literature.** *International Journal of Hydrogen Energy*, v. 26, p. 13–28, 2001.

DAVID B LEVIN, LAWRENCE PITT, MURRAY LOVE. **Biohydrogen production: prospects and limitations to practical application.** *International Journal of Hydrogen Energy*. Volume 29, Issue 2. 2004. Pages 173-185. ISSN 0360-3199[https://doi.org/10.1016/S0360-3199\(03\)00094-6](https://doi.org/10.1016/S0360-3199(03)00094-6).

DE GIOANNIS, G.; MUNTONI, A.; POLETTINI, A.; POMI, R. **A review of dark fermentative hydrogen production from biodegradable organic substrates.** *Waste Management*, v. 33, n. 6, p. 1345-1361, 2013.

DE SÁ, LÍVIAN & CAMMAROTA, MAGALI & FERREIRA-LEITAO, VIRIDIANA. (2014). **Hydrogen production by anaerobic fermentation - General aspects and possibility of using Brazilian agro-industrial wastes.** *Química Nova*. 37. 10.5935/0100-4042.20140138.

DEBABRATA DAS, T. NEJAT VEZIROGLU. **Advances in biological hydrogen production processes.** 2008. *International Journal of Hydrogen Energy*. Volume 33. Issue 21. Pages 6046-6057, ISSN 0360-3199. <https://doi.org/10.1016/j.ijhydene.2008.07.098>.

DEBABRATA DAS, T. NEJAT VEZIROĞLU. **Hydrogen production by biological processes: a survey of literature.** *International Journal of Hydrogen Energy*. Volume 26, Issue 1. Pages 13-28, ISSN 0360-3199. [https://doi.org/10.1016/S0360-3199\(00\)00058-6](https://doi.org/10.1016/S0360-3199(00)00058-6).

DOMINGOS, H. V. **Produção metanogênica a partir da codigestão anaeróbia de silagem de capim-elefante com vinhaça em reator em batelada.** 2022. Dissertação (Mestrado em Recursos Hídricos e Saneamento) – Centro de Tecnologia, Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Universidade Federal de Alagoas, Maceió, 2022.

DONG, Y.; DAI, X.; DING, Z.; FENG, H. Effect of pH on biohydrogen production from starch wastewater by an enriched anaerobic mixed microflora. *International Journal of Hydrogen Energy*, v. 34, n. 17, p. 7182-7188, 2009.

DUAN, N.; DONG, B.; WU, B.; DAI, X. **High-solids anaerobic digestion of sewage sludge under mesophilic conditions: Feasibility study.** *Environmental Science and Pollution Research*, v. 23, n. 14, p. 13418-13427, 2016.

DUTRA, J. da Cruz Ferraz; PASSOS, M. Fonseca; GARCÍA, G. Jasper Yupanqui; GOMES, R. Floripes; MAGALHÃES, T. Andrade; FREITAS, A. dos Santos; LAGUNA,

J. Guimaraes; COSTA, F. Morais Rodrigues da; SILVA, T. Fernando da; RODRIGUES, L. Santos; AMÉRICO, M. Ferrary; CAMPOS, G. Munis; PEREIRA, G.; VANDENBERGHE, L.; SOCCOL, C. Ricardo; GOÉS-NETO, A.; AZEVEDO, V. Ariston de Carvalho. **Anaerobic digestion using cocoa residues as substrate: Systematic review and meta-analysis.** *Energy for Sustainable Development*, v. 72, p. 265-277, 2023.

DUTRA, A.; FERRAZ, R.; FONSECA, L.; ERIKA, A.; MORETTI, L. **Aplicação de Pré-tratamentos em Biomassas para Aumento da Produção de Hidrogênio.** *Journal of Sustainable Bioenergy*, v. 17, n. 3, p. 320-329, 2022.

DUTRA, F. B.; FERRAZ, F. M.; FONSECA, J. F.; ERIKA, A.; MORETTI, A. S. **Optimization of biohydrogen production using agro-industrial residues in batch reactors.** *Renewable Energy*, v. 181, p. 1298-1306, 2022.

DUTRA, L. M.; *et al.* **Métodos de Tratamento de Biomassa para Produção de Biogás.** *Renewable Energy Journal*, v. 54, n. 2, p. 210-215, 2020.

EL-MASHAD, H. M.; ZHANG, R. **Biogas production from co-digestion of dairy manure and food waste.** *Bioresource Technology*, v. 101, n. 12, p. 4021-4028, 2010.

ESPOSITO, G. **A review on dark fermentative biohydrogen production from organic biomass: Process parameters and use of by-products.** *Applied Energy*, v. 144, p. 73–95, 2015.

FANG, H. H. P. *et al.* **Microbial fuel cells: A review.** *Biotechnology Advances*, v. 24, n. 2, p. 202-220, 2006.

FARAGE, *et. al.* **Avaliação do potencial de aproveitamento de aproveitamento energético dos resíduos de madeira e derivados gerados em fábricas do polo moveleiro de Ubá – MG.** *Ciência Florestal*, Santa Maria, RS, v. 23, n. 1, p. 203-205, jan/mar. 2013

FARAGE, L.; LIMA, T.; VALDEZ-VAZQUEZ, I. **Estudos sobre o reaproveitamento de resíduos para biocombustíveis**. *Bioenergia*, v. 47, p. 72-81, 2013.

FARAGE, R. C.; *et al.* **Gestão de Resíduos e Impacto Ambiental**. Estudos em Sustentabilidade Ambiental, v. 9, n. 5, p. 522-530, 2013.

FAVERO, F.; BELFIORE, P. P. **Impacto do pH, SVI e agitação na produção de hidrogênio em reatores anaeróbios**. *Revista de Bioenergia e Sustentabilidade*, v. 5, n. 1, p. 33-45, 2021.

FAVERO, F.; BELFIORE, P. P. **Normalidade dos resíduos em modelos de regressão**. *Revista de Análise Estatística e Modelagem*, v. 8, n. 2, p. 122-138, 2019.

FAVERO, F.; BELFIORE, P. P. **Significância estatística na modelagem de regressão**. *Livro Técnico de Estatística e Modelagem Aplicada*. 2017.

FAVERO, F.; *et al.* **Impacto da concentração de sólidos voláteis (SVI) e pH no rendimento de hidrogênio**. *Revista Internacional de Engenharia Bioenergética*, v. 6, n. 3, p. 157-170, 2021.

FAVERO, F.; *et al.* **Influência de variáveis operacionais e composicionais na eficiência de reatores anaeróbios**. *Revista Brasileira de Engenharia Química*, v. 10, n. 4, p. 214-227, 2022.

FAVERO, J. *et al.* **The use of Spearman correlation heatmaps in PCA: Enhancing variable selection and interpretability**. *Journal of Environmental Engineering*, v. 142, n. 1, p. 10-25, 2016.

FÁVERO, L. P. L.; BELFIORE, P. **Manual de análise de dados: estatística e modelagem multivariada com Excel, SPSS e Stata**. Rio de Janeiro: Elsevier, 2017.

GHIMIRE, A.; FRUNZO, L.; PIROZZI, F.; TRABLY, E.; ESCUDIE, R.; LENS, P. N.; ESPOSITO, G. **A review on dark fermentative biohydrogen production from organic biomass: Process parameters and use of by-products**. *Applied Energy*, v. 144, p. 73-95, 2015.

GHIMIRE, A.; SPOSITO, F.; FRUNZO, L.; TRABLY, E.; ESCUDIÉ, R.; PIROZZI, F.; LENS, P. N. L.; ESPOSITO, G. **Effects of operational parameters on dark fermentative hydrogen production from biodegradable complex waste biomass.** *Waste Management*, v. 50, p. 55-64, 2016.

GINKEL SV, SUNG S, LAY JJ. **Biohydrogen production as a function of pH and substrate concentration.** *Environ Sci Technol*. 2001 Dec 15;35(24):4726-30. doi: 10.1021/es001979r. PMID: 11775145.

GINKEL, S. V. Process parameters affecting hydrogen production by fermentation. *Water Science and Technology*, v. 44, n. 8, p. 173-178, 2001.

GOIS, G. N. S. B. *Produção Contínua de Hidrogênio em Reatores de Leito Fluidificado.* Anais da Conferência de Bioenergia, p. 45-51, 2017.

GOIS, Georgia Nayane Silva Belo. **Produção de hidrogênio e metano a partir da vinhaça de cana-de-açúcar em reatores anaeróbios de leito fluidizado.** 2017. 90 f. Dissertação (Mestrado em Engenharia Química) – Centro de Tecnologia, Programa de Pós Graduação em Engenharia Química, Universidade Federal de Alagoas, Maceió, 2017.

GHIMIRE, A., FRUNZO, L., PIROZZI, F., TRABLY, E., ESCUDIÉ, R., LENS, P. N. L., & ESPOSITO, G. (2016). **Reactors for two-stage hydrogen and methane production from sugarcane syrup.** *International Journal of Hydrogen Energy*, 41(40), 17884-17895.

GUO, X. M.; TRABLY, E.; LATRILLE, E.; CARRERE, H.; STEYER, J. P. **Hydrogen production from agricultural waste by dark fermentation: A review.** *International Journal of Hydrogen Energy*, v. 35, n. 19, p. 10660-10673, 2010.

HAIR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E. **Multivariate Data Analysis.** 7. ed. Upper Saddle River, NJ: Prentice-Hall, 2010.

HAVVA BALAT, ELIF KIRTAY. **Hydrogen from biomass – Present scenario and future prospects**. International Journal of Hydrogen Energy. 2010. Volume 35, Issue 14. Pages 7416-7426, ISSN 0360-3199. <https://doi.org/10.1016/j.ijhydene.2010.04.137>.

INTANOO, P.; KONGJAN, P.; IMAI, T. **Biohydrogen production from starch processing wastewater by anaerobic mixed cultures: Effect of initial pH**. International Journal of Hydrogen Energy, v. 39, n. 13, p. 7505-7512, 2014.

INTANOO, P.; SUTTIKUL, T.; LEETHOCHAWALIT, M.; GULARI, E.; CHAVADEJ, S. **Hydrogen production from alcohol wastewater with added fermentation residue by an anaerobic sequencing batch reactor (ASBR) under thermophilic operation**. International Journal of Hydrogen Energy, vol. 39, p 9611-9620, 2014.

J.; PRAVEENKUMAR, R. **Biohydrogen and biogas – an overview on feedstocks and enhancement process**. Fuel, v. 185, p. 810–828, 2016.

JIANG, Y. *et al.* Biohydrogen production from organic waste: An integrated approach for hydrogen generation and waste treatment. *Processes*, v. 8, n. 10, p. 1245, 2020.

JOLLIFFE, I. T.; CADIMA, J. **Principal component analysis: A review and recent developments**. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 374, n. 2065, 20150202, 2016.

KAFLE, G. K.; KIM, S. H. **Performance and kinetic modeling of anaerobic digestion of food waste**. *Bioresource Technology*, v. 144, p. 137–144, 2013.

KAISER, H. F. **The application of electronic computers to factor analysis**. *Educational and Psychological Measurement*, v. 20, n. 1, p. 141-151, 1960.

KAPARAJU, P. *et al.* **Anaerobic digestion of organic solid waste: A review**. *Waste Management*, v. 29, n. 5, p. 1502-1512, 2009.

KARADAG, D.; KÖROĞLU, O. E.; OZKAYA, B.; ÇAKMAKCI, M. **A review on anaerobic biohydrogen production from organic wastes**. *International Journal of Hydrogen Energy*, v. 39, n. 20, p. 9673-9681, 2014.

KARKI, R.; MARCHAIM, U. *et al.* **The influence of inoculum to substrate ratio on anaerobic digestion performance**. *International Journal of Renewable Energy Development*, v. 8, n. 4, p. 347-353, 2019.

KHANAL, S. K. **Production of hydrogen gas in wastewater treatment**. *Journal of Environmental Science and Health*, v. 39, n. 8, p. 2099-2107, 2004.

KOTHARI, R. *ET AL.* **Different aspects of dry anaerobic digestion for bio-energy: An overview**. *Renewable and Sustainable Energy Reviews*, v. 39, p. 174–195, 2014.

KOTHARI, R.; TYAGI, V. V.; PATHAK, A.; KUMAR, A. **Kinetic study of hydrogen production by dark fermentation using different biomass sources**. *Bioresource Technology*, v. 156, p. 45-50, 2014.

LAMAISON, C. M. **Bio-hidrogênio e resíduos sólidos: uma abordagem técnico-econômica**. Universidade Federal de Alagoas, 2009.

LAMAISON, F. C. **Aplicação da água residuária do processamento da mandioca como substrato para a produção de hidrogênio por processo fermentativo**. 2009. 83 f. Dissertação (Mestrado em Engenharia Hidráulica e Saneamento) –Universidade Federal de Santa Catarina, Florianópolis, 2009.

LAMAISON, F. C.; Reginatto, V.; Antônio, R. **Efeito de Diferentes pré-tratamentos de cultura mista para o enriquecimento em bactérias produtoras de hidrogênio**. XX Congresso Brasileiro de Engenharia Química, v. 1, n. 2, p. 717-724.

LAY, J. J. **Biohydrogen production by fermentation of starch waste with *Clostridium* species**. *Biotechnology and Bioengineering*, v. 90, n. 2, p. 185-189, 2003.

LAY, J-J.; LEE, Y-J.; NOIKE, T. **Feasibility of biological hydrogen production from organic fraction of municipal solid waste**. *Water Research*, v. 33, n. 11, p. 2579-2586, 1999.

LEE, S.; PARK, J.; OH, S. **Effects of substrate concentration and substrate-to-inoculum ratio on hydrogen production in anaerobic digestion**. *Renewable Energy*, v. 149, p. 1221-1228, 2020.

LEVIN, D. B.; PITT, L.; LOVE, M. **Biohydrogen production: Prospects and limitations to practical application**. *International Journal of Hydrogen Energy*, v. 29, n. 2, p. 173-185, 2004.

LI, Y. *et al.* **A review of carbohydrate recovery from wastewaters**. *Bioresource Technology*, v. 113, p. 191-197, 2012.

LIMA, D. M. F.. **Produção de biohidrogênio a partir de soro de leite em AnSBBR com recirculação da fase líquida**. 2015. Tese (Doutorado em Hidráulica e Saneamento) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2015. doi:10.11606/T.18.2016.tde-11052016-092510. Acesso em: 2023-05-29.

LIMA, G. P. **Produção de bio-hidrogênio a partir de frutas e vegetais residuais em reatores anaeróbios em batelada**. 2020c. 74 f. Dissertação (Mestrado em Engenharia Química) - Centro de Tecnologia, Programa de Pós-Graduação em Engenharia Química, Universidade Federal de Alagoas, Maceió, 2020.

LIMA, P. G. **Produção de Hidrogênio em Reatores Anaeróbios: Uma Revisão**. *Anais do Simpósio Brasileiro de Bioenergia*, v. 3, p. 47-55, 2020.

LIU, S. *et al.* **Enhancement of hydrogen production in dark fermentation**. *International Journal of Hydrogen Energy*, v. 34, n. 12, p. 5480-5487, 2009.

LIU, Y.; ZHANG, T.; CHEN, Y. *et al.* **Effect of inoculum volume on microbial community and hydrogen production in anaerobic digestion.** *Journal of Environmental Science and Technology*, v. 54, n. 3, p. 211-218, 2020.

LUPATINI, et. al. **Proteína microbiana em biomassa residual de reatores anaeróbios para produção de hidrogênio a partir de resíduo líquido de fecularia.** (2013). *Revista Brasileira de Energias Renováveis*. 2. 10.5380/rber.v2i2.33808.

MADEIRA, J. G. F. **Análise exergética, econômica e ambiental da produção de hidrogênio a partir da manipueira.** 2018. 145 f. Tese (Doutorado em Ciência, Tecnologia e Inovação em Agropecuária) - Pró-Reitoria de Pesquisa e Pós-Graduação, Universidade Federal Rural do Rio de Janeiro, Seropédica - RJ, 2018.

MADEIRA, L. **Hidrogênio: Um Combustível Limpo para o Futuro.** *Revista Brasileira de Energia Renovável*, v. 14, n. 3, p. 245-250, 2018.

MADEIRA, L. M. **Hydrogen as an energy vector: Properties, applications and sustainability issues.** *Renewable and Sustainable Energy Reviews*, v. 81, p. 1672-1684, 2018.

MAINTINGUER, S. I. *et al.* **Fermentative hydrogen production by microbial consortium.** *International Journal of Hydrogen Energy*, v. 33, p. 4309-4317, 2008.

MAO, C. *et al.* **Review on research achievements of biogas from anaerobic digestion.** *Renewable and Sustainable Energy Reviews*, v. 45, p. 540–555, 2015.

MAO, Y.; ZHAO, Y.; LI, S.; WANG, B. **Recent advances in biohydrogen production from lignocellulosic biomass.** *Bioresource Technology*, v. 200, p. 13-20, 2015.

MARQUES, T. D. **Produção de hidrogênio em um reator anaeróbio de leite fluidificado a partir do soro de leite.** 2019. 73 f. Dissertação (Mestrado em Recursos Hídricos e Saneamento) – Centro de Tecnologia, Programa de Pós Graduação em Recursos Hídricos e Saneamento, Universidade Federal de Alagoas, Maceió, 2018.

MARQUES, T. D. ***Produção Simultânea de Hidrogênio e Metano em Reatores Anaeróbios de Leito Fluidificado***. Revista de Energias Renováveis e Biocombustíveis, v. 8, n. 4, p. 175-183, 2018.

NETO, L. T. ***Efeito da relação C/N na produção de hidrogênio a partir da codigestão da vinhaça e capim elefante***. 2019. Dissertação (Mestrado em Recursos Hídricos e Saneamento) – Centro de Tecnologia, Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Universidade Federal de Alagoas, Maceió, 2019.

NTAIKOU, I.; ANTONOPOULOU, G.; LYBERATOS, G. ***Biohydrogen production from biomass and wastes via dark fermentation: A review***. Waste and Biomass Valorization, v. 1, n. 1, p. 21-39, 2010.

NUALSRI, C.; KONGJAN, P.; REUNGSANG, A. ***Direct integration of CSTR-UASB system for two-stage anaerobic digestion of sugarcane bagasse into hydrogen and methane production***. International Journal of Hydrogen Energy, v. 43, n. 1, p. 148-159, 2018.

NUNES, A. A. ***Produção de biohidrogênio a partir de resíduos agroindustriais***. Revista Brasileira de Ciências Agrárias, v. 10, n. 3, p. 439-445, 2015.

NUNES, A. N. ***Produção de hidrogênio a partir da manipueira em reator anaeróbio de leito fixo***. 2019. 83 f. Dissertação (Mestrado em Recursos Hídricos e Saneamento) – Centro de Tecnologia, Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Universidade Federal de Alagoas, Maceió, 2015.

OHIMAIN, E. I.; IZAH, S. C. ***A review of biogas production from palm oil mill effluents using different configurations of bioreactors***. Renewable and Sustainable Energy Reviews, v. 70, p. 242–253, 2017.

OHIMAIN, E. I.; IZAH, S. C. ***Production of hydrogen and methane from agricultural residues***. Biofuel Research Journal, v. 4, n. 1, p. 482-490, 2017.

OSTERMEYER, Y. *et al.* **Modelagem e otimização da produção de biogás a partir de resíduos orgânicos.** *Energia Renovável*, v. 129, p. 118-126, 2018.

PAGE, S.; KRUMDIECK, S. **A review of alternative fuel technologies for cars.** *Renewable and Sustainable Energy Reviews*, v. 13, n. 9, p. 2787-2793, 2009.

Page, S.; Krumdieck, S.; **"System-level energy efficiency is the greatest barrier to development of the hydrogen economy,"** 2009. *Energy Policy*, Elsevier, vol. 37(9), pages 3325-3335, September.

PEIXOTO, G. **Sistema fermentativo de duas fases para a produção de hidrogênio e metano a partir de esgoto sanitário combinado com águas residuárias industriais.** 191 f. Tese (Doutorado em Ciências: Engenharia Hidráulica e Saneamento) - Universidade de São Paulo, Escola de Engenharia de São Carlos, São Carlos, 2011.

PEIXOTO, R. S. **Microbial hydrogen production from biomass-derived sugars.** *Bioprocess and Biosystems Engineering*, v. 34, n. 5, p. 511-525, 2011.

PEREYRA, D. L. A. D. **Efeito do Tempo de Detenção Hidráulica em RALF para Produção de Hidrogênio.** *Revista de Engenharia de Recursos Naturais*, v. 10, n. 3, p. 203-211, 2019.

PEREYRA, Daniela de Los Angeles Dalmao. **Cofermentação de glicerol e melaço como alternativas para produção de biocombustíveis e produtos de valor agregado.** 2022. 87 f. Dissertação (Mestrado em Recursos Hídricos e Saneamento) - Centro de Tecnologia, Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Universidade Federal de Alagoas, Maceió, 2019.

PODDAR, S.; *et al.* **Inovações no Tratamento de Biomassa para Biocombustíveis.** *Journal of Advanced Biotechnology*, v. 12, n. 5, p. 290-298, 2021.

REGINATTO, V., & ANTÔNIO, R. V.. (2015). **Fermentative hydrogen production from agroindustrial lignocellulosic substrates. Brazilian Journal of Microbiology**, 46(2), 323–335. <https://doi.org/10.1590/S1517-838246220140111>

RENCHER, A. C. *Methods of Multivariate Analysis*. New Jersey: Wiley, 2002.

REYES, J. C.; RODRÍGUEZ, A. T.; BUITRÓN, G.; ANDRADE, I. M.; BRIONES, R. P.; FLORES, E. R.; JUÁREZ, A. G.; VARGAS, J. V.; BERNET, N.; BRAGA, A. F. M.; BRAGA, L.; CASTELLÓ, E.; CHATELLARD, L.; ETCHEBEHERE, C.; FUENTES, L.; BECERRIL, E. L.; MÉNDEZ-ACOSTA, H. O.; RUIZ-FILIPPI, G.; TAPIA-VENEGAS, E.; TRABLY, E.; WENZEL, ZAIAT, M. **A standardized biohydrogen potential protocol: An international round robin test approach**. 2019. *International Journal of Hydrogen Energy*. Volume 44, Issue 48. Pages 26237-26247. ISSN 0360-3199. <https://doi.org/10.1016/j.ijhydene.2019.08.124>.

RITTIMANN, B. E.; MCCARTY, P. L. **Environmental Biotechnology: Principles and Applications**. New York: McGraw-Hill, 2001.

Sá, L. R. V. de ., Cammarota, M. C., & Ferreira-Leitão, V. S. **Produção de hidrogênio via fermentação anaeróbia - aspectos gerais e possibilidade de utilização de resíduos agroindustriais brasileiros**. (2014) *Química Nova*, 37(5), 857–867.

Sá, L. R. V. de ., Cammarota, M. C., & Ferreira-Leitão, V. S. **Produção de hidrogênio via fermentação anaeróbia - aspectos gerais e possibilidade de utilização de resíduos agroindustriais brasileiros**. (2014). *Química Nova*, 37(5), 857–867.

SAMIR KUMAR KHANAL, WEN-HSING CHEN, LING LI, SHIHWU SUNG. **Biological hydrogen production: effects of pH and intermediate products**. 2004. *International Journal of Hydrogen Energy*. Volume 29, Issue 11. Pages 1123-1131. ISSN 0360-3199. <https://doi.org/10.1016/j.ijhydene.2003.11.002>.

SANTOS, E. C. L. **Associação de Manipueira e Esgoto Sanitário para Produção de Hidrogênio em RALF**. Revista de Processos Ambientais Sustentáveis, v. 3, n. 5, p. 55-63, 2014.

SANTOS, E.; SILVA, A. M.; NUNES, A. **Hydrogen production from sugarcane waste in anaerobic digesters**. Journal of Cleaner Production, v. 72, p. 203-210, 2014.

SHIDA, G. M.; DORNELAS, K. C.; SANTOS, A. S. **Efeito da temperatura na produção de hidrogênio por consórcio microbiano**. Revista Brasileira de Engenharia Química, v. 26, n. 2, p. 154-162, 2009.

SHOW, K. Y.; LEE, D. J.; TAY, J. H. **Biohydrogen production: Current perspectives and the way forward**. International Journal of Hydrogen Energy, v. 37, n. 18, p. 15616-15631, 2012.

SILVA, A. N. **Produção de Hidrogênio com Resíduos de Suinocultura em Reator de Leito Fluidificado**. Revista Brasileira de Bioengenharia Aplicada, v. 5, n. 2, p. 98-105, 2015.

SILVA, Olga Eliza Rocha. **Produção de biogás a partir da co-digestão anaeróbia de capim elefante com vinhaça em reator em batelada**. 2020. 87 f. Dissertação (Mestrado em Recursos Hídricos e Saneamento) – Centro de Tecnologia, Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Universidade Federal de Alagoas, Maceió, 2020.

SILVA, R. L. **Fermentação anaeróbia para a produção de biohidrogênio: Aspectos operacionais e tecnológicos**. Engenharia Sanitária e Ambiental, v. 22, n. 4, p. 589-600, 2017.

SOARES, J. D. M. **Produção de hidrogênio a partir de resíduos agroindustriais em reatores anaeróbios**. Dissertação de Mestrado, Universidade Federal de Alagoas, 2019.

SOARES, J. D. M. **Produção de Hidrogênio a partir de Resíduos de Frutas em Reatores Anaeróbios**. Simpósio Internacional de Bioengenharia, p. 87-92, 2019.

SOARES, J. D. M.. **Produção de hidrogênio a partir de cascas de banana em reator anaeróbio em batelada**. 2019. 60 f. Dissertação (Mestrado em Recursos Hídricos e Saneamento) – Centro de Tecnologia, Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Universidade Federal de Alagoas, Maceió, 2019.

STROOT, P. G.; MCMAHON, K. D.; MACKIE, R. I.; RASKIN, L. **Anaerobic codigestion of municipal solid waste and biosolids under various mixing conditions—II: Microbial population dynamics**. *Water Research*, v. 35, n. 7, p. 1804-1816, 2001.

SULAIMAN, S. A. *et al.* **Non-linear regression modeling for predicting biogas production**. *Journal of Environmental Chemical Engineering*, v. 9, n. 3, p. 105237, 2021.

TAVARES, T. G. A. **Estudo sobre a Produção de Hidrogênio em Reatores de Leito Fluidificado**. Simpósio Brasileiro de Energias Alternativas, v. 5, p. 99-107, 2016.

TAVARES, T. G. A. **Produção de hidrogênio a partir da manipueira em reator anaeróbio de leito fluidificado sob condições termofílicas**. 2022. 65 f. Dissertação (Mestrado em Recursos Hídricos e Saneamento) – Centro de Tecnologia, Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Universidade Federal de Alagoas.

TCHOBANOGLIOUS, G. *et al.* **Wastewater Engineering: Treatment and reuse**. New York: McGraw-Hill, 2003.

TELLES, N. L. S. **Efeitos da DQO no Processo de Produção de Hidrogênio**. *Revista Brasileira de Hidrogênio Sustentável*, v. 4, n. 1, p. 90-98, 2015.

TELLES, N. L. S. ***Produção de Biocombustíveis e Produtos de Valor Agregado em Reatores Anaeróbios***. Revista Brasileira de Energia Sustentável, v. 14, n. 2, p. 115-122, 2023.

TONELLO, L. ***Estudo de pré-tratamentos de resíduos lignocelulósicos para produção de hidrogênio por via anaeróbia***. Universidade Federal de São Carlos, 2017.

TONELLO, T. U. ***Produção de hidrogênio a partir da água residuária e do bagaço da extração de fécula de mandioca em AnSBBR***. 2022.104 f. Tese (Doutorado em Engenharia Agrícola) - Universidade Estadual do Oeste do Paraná, Cascavel.

V. DANDIKAS, H. HEUWINKEL, F. LICHTI, J.E. DREWES, K. KOCH. ***Correlation between biogas yield and chemical composition of energy crops***. **Bioresource Technology**. Volume 174. 2014. Pages 316-320. ISSN 0960-8524. <https://doi.org/10.1016/j.biortech.2014.10.019>.

VALDEZ-VAZQUEZ, I.; POGGI-VARALDO, H. M. ***Hydrogen production by fermentative consortia***. Renewable and Sustainable Energy Reviews, v. 13, n. 5, p. 1000-1013, 2009.

VEZIROGLU, T. N.; SAHIN, S. ***21st Century's energy: Hydrogen energy system***. **Energy Conversion and Management**, v. 49, p. 1820-1831, 2008.

VEZIROGLU, T. N.; SAHIN, S. ***21st century's energy: Hydrogen energy system***. **Energy Conversion and Management**, v. 49, n. 7, p. 1820-1831, 2008.

VEZIROGLU, T. N.; SAHIN, S. ***Impact of Hydrogen on the Environment***. International Journal of Hydrogen Energy, v. 33, n. 16, p. 409-417, 2008. Disponível em bases científicas como Scopus e Google Scholar.

VILELA, L. M. B. ***Produção biológica de hidrogênio a partir da manipueira em reator anaeróbio de leiteo fluidificado: Efeito da adição de esgoto sanitário***

para redução da carga orgânica. 2019. 77 f. Dissertação (Mestrado em Recursos Hídricos e Saneamento) – Centro de Tecnologia, Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Universidade Federal de Alagoas, Maceió, 2013.

VILELA, L. M. B. *Produção de Hidrogênio em Reator Anaeróbio de Leito Fluidificado com Manipueira.* Revista de Energias Renováveis, v. 2, n. 4, p. 38-47, 2013.

VILELA, L. M. B. **Produção biológica de hidrogênio a partir da manipueira em reator anaeróbio de leito fluidificado:** Efeito da adição de esgoto sanitário para redução da carga orgânica. 2019. 77 f. Dissertação (Mestrado em Recursos Hídricos e Saneamento) – Centro de Tecnologia, Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Universidade Federal de Alagoas, Maceió, 2013.

WANG, J. L., WAN, W. **Factors influencing fermentative hydrogen production: A review.** International Journal of Hydrogen Energy, v. 34, p. 799-811, 2009.

WANG, J.; WAN, W. **Enhancing fermentative hydrogen production using mixed microbial cultures.** International Journal of Hydrogen Energy, v. 36, n. 1, p. 277-284, 2011.

WANG, J.; WAN, W. **Factors influencing fermentative hydrogen production: A review.** International Journal of Hydrogen Energy, v. 34, n. 2, p. 799-811, 2009.

WANG, J.; WAN, W. **Factors influencing fermentative hydrogen production: a review.** Hydrogen Energy, v. 34, p. 799–811. 2009.

WHANG, L. M. **Optimization of biohydrogen production from food waste.** Environmental Science and Technology, v. 39, n. 10, p. 3889-3896, 2005.

XINMEI GUO, ERIC TRABLY, ERIC E. LATRILLE, HÉLÈNE CARRÈRE, JEAN-PHILIPPE STEYER. **Hydrogen production from agricultural waste by dark**

fermentation: A review. International Journal of Hydrogen Energy, 2010, 35 (19), pp.10660-10673. <10.1016/j.ijhydene.2010.03.008>.

YASIN, N. H. M. *et al.* **Food waste and food processing waste for biohydrogen production: a review.** Journal of Environmental Management, v. 130, p. 375–85, 2013.

YASIN, N. H. M.; LI, S.; HAMID, A.; HWAI, T. S. *Biohydrogen production by dark fermentation.* Renewable and Sustainable Energy, 2013

ZHANG, Y. *et al.* Effect of temperature on anaerobic digestion of organic waste. *Environmental Science & Technology*, v. 47, n. 2, p. 565-574, 2013.

ZHOU, Y. *et al.* Continuous biohydrogen production from mixed wastewater. *International Journal of Hydrogen Energy*, v. 35, n. 21, p. 11778-11787, 2010.