



Michel Angelo Oliveira Werneck de Carvalho

**Avaliação de modelos preditivos para produção de biogás
usando redes neurais artificiais**

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-graduação em Engenharia Mecânica da PUC-Rio como requisito parcial para obtenção do grau de Mestre em Engenharia Mecânica.

Orientador: Prof. Florian Alain Yannick Pradelle

Co-orientador: Prof. Brunno Ferreira dos Santos

Rio de Janeiro

Abril de 2023



Michel Angelo Oliveira Werneck de Carvalho

**Avaliação de modelos preditivos para produção de biogás
usando redes neurais artificiais**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Pós-graduação em Engenharia Mecânica da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Florian Alain Yannick Pradelle
Orientador
Departamento de Engenharia Mecânica – PUC-Rio

Prof. Brunno Ferreira dos Santos
Coorientador
Departamento de Engenharia Química e Materiais – PUC-Rio

Prof. Helon Vicente Hultmann Ayala
Departamento de Engenharia Mecânica – PUC-Rio

Profa. Amanda Lemette Teixeira Brandão
Departamento de Engenharia Química e de Materiais – PUC-Rio

Profa. Sibeles Augusta Ferreira Leite
Universidade Federal de Viçosa – UFV

Rio de Janeiro, 27 de abril de 2023

Todos os direitos reservados. A reprodução, total ou parcial, do trabalho é proibida sem autorização do autor, do orientador e da universidade.

Michel Angelo Oliveira Werneck de Carvalho

Graduou-se em Engenharia Mecânica pela Pontifícia Universidade Católica do Rio de Janeiro (2018) e em Engenharia de Petróleo pela Pontifícia Universidade Católica do Rio de Janeiro (2014). Tem experiência na área de Engenharia Mecânica.

Ficha Catalográfica

Carvalho, Michel Angelo Oliveira Werneck de

Avaliação de modelos preditivos para produção de biogás usando redes neurais artificiais / Michel Angelo Oliveira Werneck de Carvalho ; orientador: Florian Alain Yannick Pradelle ; co-orientador: Brunno Ferreira dos Santos. – 2023.

154 f. : il. color. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Mecânica, 2023.

Inclui bibliografia

1. Engenharia Mecânica – Teses. 2. Metano. 3. Resíduos de frutas e vegetais. 4. Co-digestão. 5. Resíduos alimentares. 6. Biodigestão. I. Pradelle, Florian Alain Yannick. II. Santos, Brunno Ferreira dos. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Mecânica. IV. Título.

CDD: 621

Agradecimentos

Gostaria de expressar minha sincera gratidão a todas as pessoas que contribuíram para o sucesso desta dissertação de mestrado.

Agradeço primeiramente ao meu orientador, Florian, pela orientação excepcional, paciência e apoio constante ao longo deste processo. Seu conhecimento e perspicácia foram fundamentais para o desenvolvimento deste trabalho. Sua orientação não apenas aprimorou a qualidade acadêmica deste estudo, mas também contribuiu significativamente para o meu crescimento como pesquisador. Este trabalho não teria alcançado o nível de sucesso sem a sua valiosa contribuição.

Ao meu coorientador, Bruno, igualmente agradeço pela orientação adicional, contribuições valiosas e pelo suporte contínuo. Sua competência foi essencial para a conclusão desta pesquisa.

Aos membros da banca examinadora, Helon, Amanda e Sibebe, agradeço por dedicarem seu tempo e sabedoria na avaliação deste trabalho. Suas sugestões e observações foram extremamente construtivas.

À minha família, expressei minha profunda gratidão pelo apoio incondicional, compreensão e incentivo ao longo dos anos. Suas palavras de encorajamento foram uma fonte constante de motivação.

À PUC-Rio, pela oportunidade de realizar este mestrado, agradeço pela infraestrutura e recursos disponibilizados que foram essenciais para a conclusão deste trabalho.

Por fim, a todos que, de alguma forma, contribuíram para este projeto, meu muito obrigado. Este trabalho não teria sido possível sem a ajuda e suporte de cada um de vocês.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

De Carvalho, Michel Angelo Oliveira Werneck; Pradelle, Florian Alain Yannick (Orientador); dos Santos, Brunno Ferreira (Coorientador). **Avaliação de modelos preditivos para produção de biogás usando redes neurais artificiais**. Rio de Janeiro, 2023. 154p. Dissertação de Mestrado - Departamento de Engenharia Mecânica, Pontifícia Universidade Católica do Rio de Janeiro.

O biogás é uma energia renovável com grande potencial de produção a partir de resíduos, incluindo resíduos alimentares. Nesse contexto, o presente trabalho apresenta o desenvolvimento de três modelos distintos usando Redes Neurais Artificiais (RNAs), com a capacidade de prever o volume acumulado de biogás, de metano e a concentração de CH₄, respectivamente. Foi construído um banco de dados da literatura com variáveis do processo de biodigestão anaeróbia: tipo de biomassa, tipo de reator/alimentação, teor de sólido volátil, pH, taxa de carga orgânica, tempo de retenção hidráulica, temperatura e volume do reator. Para cada conjunto de modelos, foram desenvolvidas e testadas 24 RNAs utilizando a ferramenta computacional MATLAB. As RNAs foram avaliadas pela sua capacidade de estimação através do coeficiente de determinação (R^2) e também através da soma do erro quadrático (SSE) obtidos. Após as etapas iniciais, as redes neurais foram usadas para criar superfícies de resposta, buscando regiões ideais para produção de biogás e metano. Contudo, um único modelo não atingiu a representatividade desejada, levando à segmentação dos dados por tipo de biomassa. As RNAs desenvolvidas demonstraram eficácia na estimação dos grupos usados para treinamento, teste e validação. A melhor rede alcançou R^2 de 0,9969 para biogás, 0,9963 para metano e 0,9386 para a porcentagem de metano, com SSE de 0,1808, 0,1089 e 11,45, respectivamente. A estratégia de combinar variáveis do processo em superfícies de resposta revelou-se útil para identificar pontos ótimos no processo produtivo.

Palavras-chaves

Metano; Resíduos de frutas e vegetais; Co-digestão; Resíduos alimentares; Biodigestão; Otimização

Abstract

De Carvalho, Michel Angelo Oliveira Werneck; Pradelle, Florian Alain Yannick (Advisor); dos Santos, Brunno Ferreira (Co-advisor). **Assessment of predictive models for biogas production using artificial neural networks**. Rio de Janeiro, 2023. 154p. Dissertação de Mestrado - Departamento de Engenharia Mecânica, Pontifícia Universidade Católica do Rio de Janeiro.

Biogas is a renewable energy source with significant production potential from various waste materials, including food waste. In this context, this study presents the development of three distinct models using Artificial Neural Networks (ANNs), capable of predicting the cumulative volume of biogas, methane, and CH₄ concentration, respectively. A literature-based database was constructed, including variables from anaerobic digestion processes: biomass type, reactor/feed type, volatile solid content, pH, organic loading rate, hydraulic retention time, temperature, and reactor volume. For each set of models, 24 ANNs were developed and tested using the MATLAB computational tool. The ANNs' estimation capability was assessed using the coefficient of determination (R²) and the sum of squared errors (SSE). Following initial stages, neural networks were employed to create response surfaces, aiming to identify optimal regions for biogas and methane production. However, a single model failed to achieve the desired representativeness, leading to data segmentation based on biomass type. The developed ANNs demonstrated effectiveness in estimating the groups used for training, testing, and validation. The best network achieved R² values of 0.9969 for biogas, 0.9963 for methane, and 0.9386 for methane percentage, with SSE values of 0.1808, 0.1089, and 11.45, respectively. The strategy of combining process variables in response surfaces proved valuable in identifying optimal points in the production process.

Keywords

Methane; Fruit and vegetable waste; Co-digestion; Food waste; Biodigestion; Optimization

Sumário

1	Introdução.....	18
1.1	Contextualização.....	18
1.2	Objetivos.....	24
1.3	Estrutura do trabalho.....	25
2	Revisão bibliográfica.....	26
2.1	Processo de formação do biogás.....	26
2.2	Fatores que influenciam a produção de biogás.....	28
2.3	Características e purificação do biogás.....	37
2.4	Mercado do biogás no Brasil.....	39
2.5	Fundamentos sobre redes neurais artificiais (RNAs).....	45
2.5.1	Arquiteturas das RNA's.....	46
2.5.2	Treinamento das RNA's.....	49
2.5.3	Métricas de avaliação de desempenho.....	53
2.6	Uso de redes neurais artificiais na temática do uso do biogás.....	54
3	Metodologia.....	61
3.1	Definição da base de dados com todas as biomassas e refinamento das amostras.....	62
3.2	Definição do modelo aplicado a estimação e treinamento, validação e teste da RNA.....	66
3.3	Geração das superfícies de resposta.....	68
3.4	Desenvolvimento e avaliação de modelos de RNA específicos por biomassa.....	69
4	Resultados e Discussões.....	71
4.1	Consolidação dos bancos de dados completos.....	71
4.2	Treinamento, validação e teste das redes dos bancos completos.....	80
4.2.1	Produção acumulada de biogás.....	81
4.2.2	Produção acumulada de metano.....	84
4.2.3	Concentração volumétrica de metano (CH ₄) no biogás.....	87
4.3	Superfícies de resposta para bancos de dados completos.....	90
4.3.1	Superfícies de resposta para a produção acumulada de biogás.....	90
4.3.2	Superfícies de resposta para produção de metano.....	96
4.3.3	Superfícies de resposta para concentração volumétrica de metano (CH ₄) no biogás.....	102

4.3.4	Avaliação dos modelos de produção de biogás e metano com o conjunto de dados completo	111
4.4	Resultados dos modelos específicos por biomassa e comparações	112
4.4.1	PAB - Banco de dados, modelos e superfícies.....	112
4.4.2	PAM - Banco de dados, modelos e superfícies.....	120
4.4.3	%CH ₄ - Banco de dados, modelos e superfícies	127
4.4.4	Avaliação dos modelos de produção de biogás e metano com o conjunto de dados específico por biomassa	135
5	Conclusão e sugestões para trabalhos futuros	136
6	Referências bibliográficas	141

Lista de Figuras

Figura 1 – Participação de cada caminho tecnológico na redução das emissões até 2050. Fonte: Elaboração própria. Adaptação IRENA, 2021.	19
Figura 2 - Etapas metabólicas do processo de digestão anaeróbia (Adaptado de Oliveira e Ramalho, 2006).....	27
Figura 3 - Crescimento do número de novas plantas de biogás no Brasil na última década. Fonte: Elaboração própria. Adaptado de CIBiogas 2021	43
Figura 4 - Modelo artificial do neurônio. Adaptado de Araujo, 2020.....	45
Figura 5 - Exemplo de funcionamento do algoritmo Backpropagation. Fonte: Elaboração própria. Adaptado de Lee et al., 2020	51
Figura 6 - Diagrama com as etapas da metodologia. Fonte: Elaborada pelo autor	61
Figura 7 - Histogramas para os bancos de dados PAB: comparação das distribuições de “PAB sem refinamento” e “PAB refinado” para tipo de biomassa, tipo de reator, teor de sólidos voláteis (VS), pH, taxa de carga orgânica (OLR), tempo de retenção hidráulica (HRT), temperatura, volume do reator e produção acumulada de biogás	73
Figura 8 - Histogramas para os bancos de dados PAM: comparação das distribuições de “PAM sem refinamento” e “PAM refinado” para tipo de biomassa, tipo de reator, teor de sólidos voláteis (VS), pH, taxa de carga orgânica (OLR), tempo de retenção hidráulica (HRT), temperatura, volume do reator e produção acumulada de metano	76
Figura 9 - Histogramas para os bancos de dados %CH ₄ : comparação das distribuições de “%CH ₄ sem refinamento” e “%CH ₄ refinado” para tipo de biomassa, tipo de reator, teor de sólidos voláteis (VS), pH, taxa de carga orgânica (OLR), tempo de retenção hidráulica (HRT), temperatura, volume do reator e concentração volumétrica de metano	79
Figura 10 - Resultados para treinamento, teste e validação da RNA. Fonte: Elaboração própria.....	83
Figura 11 Resultado do treinamento, teste e validação da RNA. Fonte: Elaboração própria.....	86
Figura 12 - Resultado do treinamento, teste e validação da RNA para %CH ₄ . Fonte: Elaboração própria.....	89

Figura 13 - Superfície de resposta para a combinação entre pH e volume de reator para CD.....	92
Figura 14 - Superfície de resposta para a combinação entre pH e volume de reator para FW	92
Figura 15 - Superfície de resposta para a combinação entre pH e volume de reator para FVW	93
Figura 16 - Superfície de resposta para a combinação entre temperatura e volume de reator para CD.....	94
Figura 17 - Superfície de resposta para a combinação entre temperatura e volume de reator para FW	94
Figura 18 - Superfície de resposta para a combinação entre temperatura e volume de reator para FVW	94
Figura 19 - Superfície de resposta para a combinação entre VS e volume de reator para CD.....	95
Figura 20 - Superfície de resposta para a combinação entre VS e volume de reator para FW	95
Figura 21 - Superfície de resposta para a combinação entre VS e volume de reator para FVW	96
Figura 22 - Superfície de resposta para a combinação entre pH x volume de reator para CD, PAM.....	98
Figura 23 - Superfície de resposta para a combinação entre pH x volume de reator para FW, PAM.....	99
Figura 24 - Superfície de resposta para a combinação entre pH x volume de reator para FVW, PAM.....	99
Figura 25 - Superfície de resposta para a combinação entre VS x volume de reator para CD, PAM.....	100
Figura 26 - superfície de resposta para a combinação entre VS x volume de reator para FW, PAM.....	100
Figura 27 - Superfície de resposta para a combinação entre VS x volume de reator para FVW, PAM.....	101
Figura 28 - Superfície de resposta para configuração de reator x pH para CD. B1 (Batelada 1 estágio = -3), B2 (Batelada 2 estágios = -2), SC1 (Semi contínuo 1	

estágio = -1), SC2 (Semi contínuo 2 estágios = 0), SC3 (Semi contínuo 3 estágios= 1), C2 (Contínuo 2 estágios = 2), C1 (Contínuo 1 estágio = 3).....	103
Figura 29 - Superfície de resposta para configuração de reator x pH para FW. B1 (Batelada 1 estágio = -3), B2 (Batelada 2 estágios = -2), SC1 (Semi contínuo 1 estágio = -1), SC2 (Semi contínuo 2 estágios = 0), SC3 (Semi contínuo 3 estágios= 1), C2 (Contínuo 2 estágios = 2), C1 (Contínuo 1 estágio = 3).....	104
Figura 30 - Superfície de resposta para configuração de reator x pH para FVW. B1 (Batelada 1 estágio = -3), B2 (Batelada 2 estágios = -2), SC1 (Semi contínuo 1 estágio = -1), SC2 (Semi contínuo 2 estágios = 0), SC3 (Semi contínuo 3 estágios= 1), C2 (Contínuo 2 estágios = 2), C1 (Contínuo 1 estágio = 3).....	104
Figura 31 - Superfície de resposta para pH x volume de reator para CD.....	105
Figura 32 - Superfície de resposta para pH x volume de reator para FW.	106
Figura 33 - Superfície de resposta para pH x volume de reator para FVW.	106
Figura 34 - Superfície de resposta para OLR x volume de reator para CD, %CH4.	107
Figura 35 - Superfície de resposta para OLR x volume de reator para FW, %CH4.	107
Figura 36 - Superfície de resposta para OLR x volume de reator para FVW, %CH4.	108
Figura 37 - Superfície de resposta para temperatura x volume de reator para CD, %CH4.	109
Figura 38 - Superfície de resposta para temperatura x volume de reator para FW, %CH4.	109
Figura 39 - Superfície de resposta para temperatura x volume de reator para FVW, %CH4.	110
Figura 40 - Gráfico para tipo de substrato x pH, %CH4.	111
Figura 41 - Histogramas para os bancos de dados PAB por substrato: figuras das distribuições de “PAB substrato FVW” e “PAB substrato FW” para tipo de biomassa, tipo de reator, teor de sólidos voláteis (VS), pH, taxa de carga orgânica (OLR), tempo de retenção hidráulica (HRT), temperatura, volume do reator e produção acumulada de biogás.....	114

Figura 42 – a) Resultados para treinamento, teste e validação da RNA PAB substrato FVW, b) Resultados para treinamento, teste e validação da RNA PAB substrato FW. Fonte: Elaboração própria.	116
Figura 43 – Comparação entre as superfícies de resposta para a combinação entre pH x volume de reator para PAB. A) Superfície usando banco PAB substrato FVW, B) Superfície usando banco PAB refinado completo, C) Superfície usando banco PAB substrato FW, D) Superfície usando banco PAB refinado completo.	117
Figura 44 - Comparação entre as superfícies de resposta para a combinação entre temperatura x volume de reator para PAB. A) Superfície usando banco PAB substrato FVW, B) Superfície usando banco PAB refinado completo, C) Superfície usando banco PAB substrato FW, D) Superfície usando banco PAB refinado completo.	118
Figura 45 - Comparação entre as superfícies de resposta para a combinação entre VS x Volume de reator para PAB. A) Superfície usando banco PAB substrato FVW, B) Superfície usando banco PAB refinado completo, C) Superfície usando banco PAB substrato FW, D) Superfície usando banco PAB refinado completo.	119
Figura 46 - Histogramas para os bancos de dados PAM por substrato: figuras das distribuições de “PAM substrato FVW” e “PAM substrato FW” para tipo de biomassa, tipo de reator, teor de sólidos voláteis (VS), pH, taxa de carga orgânica (OLR), tempo de retenção hidráulica (HRT), temperatura, volume do reator e produção acumulada de biogás.	122
Figura 47 - a) Resultados para treinamento, teste e validação da RNA PAM substrato FVW, b) Resultados para treinamento, teste e validação da RNA PAM substrato FW. Fonte: Elaboração própria.	124
Figura 48 - Comparação entre as superfícies de resposta para a combinação entre pH x volume de reator para PAM. A) Superfície usando banco PAM substrato FVW, B) Superfície usando banco PAM refinado completo, C) Superfície usando banco PAM substrato FW, D) Superfície usando banco PAM refinado completo.	125
Figura 49 - Comparação entre as superfícies de resposta para a combinação entre VS x volume de reator para PAM. A) Superfície usando banco PAM substrato FVW, B) Superfície usando banco PAM refinado completo, C) Superfície usando	

banco PAM substrato FW, D) Superfície usando banco PAM refinado completo.	126
Figura 50 - Histogramas para os bancos de dados %CH ₄ por substrato: figuras das distribuições de “%CH ₄ substrato FVW” e “%CH ₄ substrato FW” para tipo de biomassa, tipo de reator, teor de sólidos voláteis (VS), pH, taxa de carga orgânica (OLR), tempo de retenção hidráulica (HRT), temperatura, volume do reator e produção acumulada de biogás.....	129
Figura 51 - a) Resultados para treinamento, teste e validação da RNA %CH ₄ substrato FVW, b) Resultados para treinamento, teste e validação da RNA %CH ₄ substrato FW. Fonte: Elaboração própria.	131
Figura 52 - Comparação entre as superfícies de resposta para a combinação entre pH x volume de reator para %CH ₄ . A) Superfície usando banco %CH ₄ substrato FVW, B) Superfície usando banco %CH ₄ refinado completo, C) Superfície usando banco %CH ₄ substrato FW, D) Superfície usando banco %CH ₄ refinado completo.	132
Figura 53 - Comparação entre as superfícies de resposta para a combinação entre OLR x volume de reator para %CH ₄ . A) Superfície usando banco %CH ₄ substrato FVW, B) Superfície usando banco %CH ₄ refinado completo, C) Superfície usando banco %CH ₄ substrato FW, D) Superfície usando banco %CH ₄ refinado completo.	133
Figura 54 - Comparação entre as superfícies de resposta para a combinação entre temperatura x volume de reator para %CH ₄ . A) Superfície usando banco %CH ₄ substrato FVW, B) Superfície usando banco %CH ₄ refinado completo, C) Superfície usando banco %CH ₄ substrato FW, D) Superfície usando banco %CH ₄ refinado completo.	134

Lista de tabelas

Tabela 1 - Biomassas geralmente utilizadas em processos de digestão anaeróbia e suas características principais. Fonte: Adaptação Vasco Correa (2018).	22
Tabela 2 - Possíveis percentuais dos constituintes do biogás. Fonte: Adaptado de Cassini (2003).	37
Tabela 3 - Características dos gases presentes no biogás. Fonte: Adaptado de Gandolla et al., 1997.	38
Tabela 4 - Necessidade de remoção de compostos gasosos de acordo com a utilização do biogás. Fonte: Adaptado de Littig (2011).	39
Tabela 5 - Revisão da literatura de artigos que tratam de RNA aplicada à produção de biogás e biometano (continua).....	56
Tabela 6 - Propriedades selecionadas no levantamento de dados. Fonte: Elaboração própria.....	63
Tabela 7 - Tabela com as variáveis discretas. Fonte: Elaboração própria.....	65
Tabela 8 - Condições para definição do modelo da RNA. Fonte: Elaboração própria	67
Tabela 9 - Topologias das RNAs para produção acumulada de biogás. Fonte: Elaboração própria.....	82
Tabela 10 - Topologias das RNAs para produção acumulada de metano. Fonte: Elaboração própria.....	85
Tabela 11 - Topologias das RNAs para concentração volumétrica de CH ₄ no biogás. Fonte: Elaboração própria	88
Tabela 12 - Valores fixos das variáveis. Fonte: Elaboração própria	91
Tabela 13 - Valores de máximo e mínimo das variáveis para o banco de dados PAB treinamento. Fonte: Elaboração própria	91
Tabela 14 - Valores fixos das variáveis. Fonte: Elaboração própria	97
Tabela 15 - Valores de máximo e mínimo das variáveis para o banco de dados PAM treinamento. Fonte: Elaboração própria	98
Tabela 16 - Valores fixos das variáveis. %CH ₄ . Fonte: Elaboração própria	102
Tabela 17 - Valores de máximo e mínimo das variáveis para o banco de dados %CH ₄ treinamento. Fonte: Elaboração própria	103

Tabela 18 - Quantidade de amostras em cada subconjunto de dados PAB. Fonte: Elaboração própria.....	112
Tabela 19 - Comparação entre as configurações das melhores redes PAB. Fonte: Elaboração própria.....	116
Tabela 20 - Quantidade de amostras em cada subconjunto de dados PAM. Fonte: Elaboração própria.....	120
Tabela 21 - Comparação entre as configurações das melhores redes PAM. Fonte: Elaboração própria.....	123
Tabela 22 - Quantidade de amostras em cada subconjunto de dados %CH ₄ . Fonte: Elaboração própria.....	127
Tabela 23 - Comparação entre as configurações das melhores redes %CH ₄ . Fonte: Elaboração própria.....	130

Abreviaturas e siglas

%CH ₄	Porcentagem de metano
ABIOGAS	Associação Brasileira de Biogás
ABIQUIM	Associação Brasileira da Indústria Química
ABRELPE	Associação Brasileira de Empresas de Limpeza Pública e Resíduos Especiais
ANEEL	Agência Nacional de Energia Elétrica
ANN	<i>Artificial neural network</i> (Rede Neural Artificial)
ANP	Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
API	<i>Application Programming Interface</i> (Interface de Programação de Aplicação)
B	Reator Batelada
BNDES	Banco Nacional de Desenvolvimento Econômico e Social
C	Reator Contínuo
CD	Codigestão
CH ₃ COOH	Ácido acético
CH ₄	Metano
CO ₂	Dióxido de carbono
FWW	<i>Fruit and vegetable waste</i> (Resíduos de frutas e vegetais)
FW	<i>Food Waste</i> (Resíduos alimentares)
GEE	Gases do efeito estufa
H ₂	Hidrogênio
H ₂ S	Sulfeto de hidrogênio
HRT	<i>Hydraulic retention time</i> (Tempo de retenção hidráulica)
IEA	<i>International Energy Agency</i> (Agência Internacional de Energia)
IPNI	<i>International Plant Names Index</i> (Índice Internacional de Nomes de Plantas)
MLP	<i>Multi Layers Perceptron</i> (Perceptron multicamadas)
OLR	Organic loading rate (Taxa de carregamento orgânico)
PAB	Produção acumulada de biogás
PAM	Produção acumulada de metano
pH	Potencial Hidrogeniônico

RBF	Radial basis function
REN21	<i>Renewable Energy Policy Network for the 21st Century</i>
RNA	Rede Neural Artificial
RSU	Resíduos Sólidos Urbanos
SC	Reator Semi Contínuo
SSE	Sum of squared errors (Soma dos erros quadrados)
TS	<i>Total solids</i> (Sólidos Totais)
VS	<i>Volatile solids</i> (Sólidos Voláteis)
WBA	World Bioenergy Association
WtE	Waste-to-Energy

1 Introdução

1.1 Contextualização

O grande desafio global de combate à mudança climática está levando à adoção crescente de fontes de energia renovável, uma vez que o predomínio dos combustíveis fósseis enfrenta cada vez mais obstáculos como a volatilidade de preços e a tendência de médio e longo prazos de diminuição na oferta (Milanez et al., 2018). Apesar disso, a matriz energética mundial ainda é fortemente dependente de fontes fósseis, como o petróleo (31 %) e o carvão (27 %), enquanto as fontes renováveis de energia representam somente 14 % da produção total de energia (WBA, 2021). Essa transição energética é necessária para enfrentar os desafios das mudanças climáticas, da segurança energética e do desenvolvimento socioeconômico.

O relatório da Agência Internacional para as Energias Renováveis (IRENA) de 2021 mostra que as energias renováveis podem fornecer 90% da energia total até 2050, com uma economia líquida de US\$ 164 trilhões. Para isso, são necessárias medidas como a eletrificação dos setores, a melhoria da eficiência energética, a cooperação internacional e a inovação tecnológica (IRENA, 2021). Uma das alternativas para a transição energética é a descarbonização, que consiste em reduzir as emissões de gases de efeito estufa provenientes da produção e do consumo de energia. As fontes renováveis, como a solar, a eólica, a hidrelétrica, a biomassa e a geotérmica, podem contribuir para a descarbonização, pois não emitem ou emitem muito pouco carbono na geração de eletricidade, calor ou combustíveis. A Figura 1 abaixo mostra o potencial de redução das emissões de CO₂ até 2050 através de seis caminhos tecnológicos diferentes.

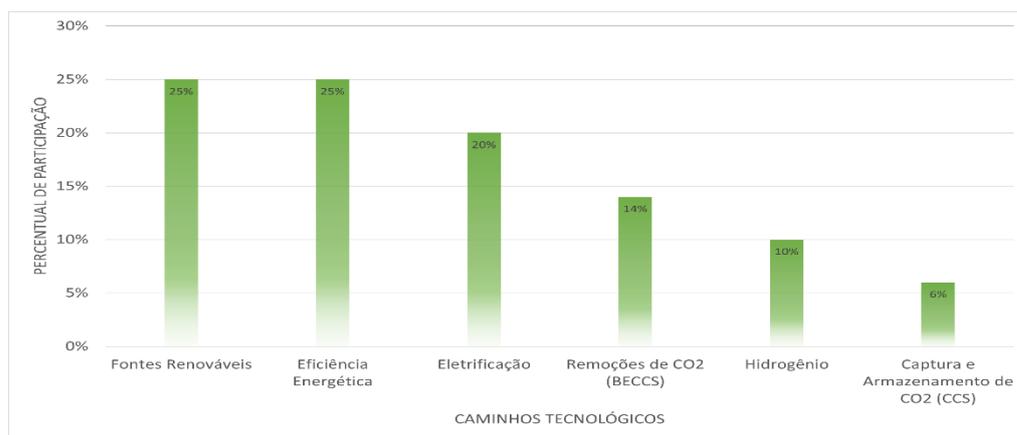


Figura 1 – Participação de cada caminho tecnológico na redução das emissões até 2050. Fonte: Elaboração própria. Adaptação IRENA, 2021.

A bioenergia, que é a energia obtida a partir da biomassa, possui características bastante particulares que a tornam um pilar de qualquer sistema sustentável de energia, tanto por sua importância atual como fonte renovável de combustível quanto por seu potencial de crescimento para as próximas décadas. Segundo o relatório REN21 de 2022, a bioenergia é a maior fonte de energia renovável no mundo, representando 12,6% do consumo energético global em 2020. No entanto, o relatório também aponta que a bioenergia enfrenta vários desafios para aumentar sua participação nos setores de uso final, como a concorrência pelo uso da terra, a sustentabilidade ambiental e social, as barreiras regulatórias e financeiras e a falta de políticas e incentivos adequados. O relatório defende que a bioenergia moderna e sustentável pode desempenhar um papel importante na mitigação das mudanças climáticas, reduzindo as emissões de CO₂ e substituindo os combustíveis fósseis, desde que seja utilizada de forma eficiente e integrada com outras fontes renováveis. Além da diversidade de fontes da qual deriva a biomassa, existe uma ampla diversidade de produtos combustíveis dela derivados e que podem produzir energia, como o bioetanol, o biodiesel e o biogás (Theuretzbacher et al., 2013; Serrano-Ruiz, 2020). A biomassa pode ser transformada através de vários processos de conversão em biocombustíveis sólidos, líquidos ou gasosos que, por sua vez, podem produzir calor, eletricidade e combustíveis para os meios de transporte (Santos, 2009).

Segundo a Agência Internacional de Energia Renovável (IRENA, 2021), a capacidade mundial de geração de energia elétrica com biogás passou de 2.455 MW em 2000 para 20.150 MW em 2020, representando um crescimento de mais de 700%. Sete países são responsáveis por 73,8% dessa capacidade instalada: Alemanha (7.459 MW), Estados Unidos (2.291 MW), Reino Unido (1.858 MW), Itália (1.432 MW), China (903 MW), França (511 MW) e Brasil (417 MW). A produção e o uso do biogás variam de acordo com as características e as políticas de cada região e país. A Europa e a Ásia são as regiões que mais se destacam na geração de biogás, respondendo por 85% de toda a produção comercializada no mundo, segundo a Associação Mundial do Biogás (WBA, 2021). Os Estados Unidos também estão liderando o caminho globalmente no uso de biometano no setor de transporte (IEA, 2020). O biometano é considerado um combustível avançado, que pode reduzir as emissões de gases de efeito estufa em até 90% em comparação com os combustíveis fósseis (Carvalho et al., 2019). Portanto, tendo em vista a direção seguida pelas nações ao redor do mundo e a crise ambiental vigente, é necessário que as pesquisas nacionais também se debruçam sobre utilização da biomassa na produção de energia (Camargo, 2012), em especial à produção de biogás, visto que, segundo Coelho (2018), o Brasil apresenta elevada capacidade e potencial de produção deste produto. De acordo com a Abiogás, o Brasil produziu cerca de 2,3 bilhões de metros cúbicos de biogás por ano em 2021, aproveitando apenas 2,7% do seu potencial. A entidade estima que o país pode chegar a produzir 30 milhões de metros cúbicos por dia em 2030, com investimentos de R\$ 60 bilhões em novas usinas. Desse modo, a digestão anaeróbica surge como uma tecnologia viável para o tratamento dos resíduos orgânicos, transformando-os em biogás e reduzindo a poluição ambiental (Silva et al., 2020). O biogás produzido pode ser utilizado como fonte de energia de forma acessível e com baixa tecnologia (Sampaio et al., 2021; Souza et al., 2022). Sua queima contribui para a redução do impacto climático negativo, tornando a atividade de processamento de resíduos lucrativa (Milanez et al., 2018). Com processamento adicional para purificação, o biogás pode substituir o gás natural em diversas aplicações (Milanez et al., 2018).

O aumento populacional global e o crescimento econômico têm apresentado desafios significativos na gestão dos Resíduos Sólidos Urbanos (RSU) nos últimos

anos (Milanez et al., 2018). A utilização dos RSU como fonte de energia por meio das rotas de conversão chamadas Waste-to-Energy (WtE) tem surgido como uma alternativa promissora, alinhada à economia circular e contribuindo para a redução, reutilização e melhoria da destinação final dos resíduos (Silva et al., 2020; Sampaio et al., 2021; Souza et al., 2022).

O Brasil, como o 4º maior produtor de resíduos sólidos do mundo (ABREN), enfrenta desafios consideráveis nesse setor. Com uma população de cerca de 208 milhões de habitantes, a maioria residindo em áreas urbanas, o país gera anualmente aproximadamente 80 milhões de toneladas de RSU. Desse total, 17,2% são destinados a lixões, 24,1% a aterros controlados e 58,7% a aterros sanitários, que são os métodos mais adequados para esse fim (ABRELPE, 2019). A aprovação do Plano Nacional de Resíduos Sólidos (Planares) pelo Decreto nº 11.043, em abril de 2022, estabelece metas, diretrizes, estratégias e ações para a gestão dos RSU, enfatizando a eliminação dos lixões até 2024 (Ecogen, 2022). O aproveitamento energético dos RSU é uma abordagem promissora e viável para a geração de energia limpa e sustentável no Brasil. A utilização do biogás produzido pela digestão anaeróbica dos resíduos orgânicos pode contribuir significativamente para a redução do impacto ambiental e dos custos associados ao tratamento dos resíduos urbanos. É fundamental investir em tecnologias e políticas que promovam o aproveitamento eficiente dos RSU, visando a sustentabilidade ambiental e energética do país.

Um aspecto importante da gestão dos RSU é o tratamento dos resíduos orgânicos, especialmente os provenientes de alimentos. Esses resíduos podem ser classificados em resíduos alimentares (Food Waste - FW), que são os alimentos descartados por motivos diversos, e resíduos de frutas e vegetais (Fruit and Vegetable Waste - FVW), que são os restos de cascas, sementes e partes não comestíveis desses alimentos. A gestão adequada desses resíduos é fundamental para a sustentabilidade da gestão de resíduos (Milanez et al., 2018). Neste trabalho, o foco será nos resíduos alimentares (FW) e nos resíduos de frutas e vegetais (FVW). Segundo a Organização das Nações Unidas para Agricultura e Alimentação (FAO) apesar dos esforços para reduzir o desperdício de alimentos, cerca de 931 milhões de toneladas de alimentos foram perdidas ou desperdiçadas em 2019. A redução dessas perdas pode trazer benefícios econômicos, sociais e ambientais

significativos, como o aumento da disponibilidade de alimentos e a redução da emissão de gases de efeito estufa. Além disso, a utilização desses resíduos como matéria-prima para a produção de biogás pode ser uma alternativa sustentável para a geração de energia.

Os resíduos de frutas e vegetais (FVW) e os resíduos alimentares (FW) são ricos em matéria orgânica, mas também podem causar desequilíbrios no processo de digestão anaeróbica. Para evitar isso, é preciso misturar esses resíduos com outros tipos de resíduos ou adicionar substâncias que neutralizem a acidez. Uma boa alternativa é fazer a codigestão entre FW e FVW, que consiste em digerir os dois resíduos juntos. Isso pode aumentar a produção de biogás e a estabilidade do processo (Shen et al., 2013). A codigestão também pode melhorar o equilíbrio entre carbono e nitrogênio, que é um fator importante para o crescimento dos microrganismos anaeróbios (El-Mashad e Zhang, 2010). Além disso, a codigestão pode aproveitar o potencial de diferentes resíduos orgânicos, reduzindo o seu impacto ambiental e gerando energia renovável (Zarkadas et al., 2015).

A Tabela 1 a seguir mostra os diferentes tipos de matéria orgânica que podem ser aplicados ao processo de digestão anaeróbica, com o seu potencial de produção de biogás em $\text{m}^3 / \text{kg VS}$.

Tabela 1 - Biomassas geralmente utilizadas em processos de digestão anaeróbia e suas características principais. Fonte: Adaptação Vasco Correa (2018).

Tipo de substrato	Características	TS [%]	Produção de biogás [$\text{m}^3/\text{kg VS}$]
Dejetos de animais	-Geralmente codigerido com biomassas com alto teor de carbono	2 - 20	0,1 – 0,6
Fração orgânica de resíduos sólidos urbanos (RSU)	- Necessita de separação na fonte, ou alto pré-processamento. - Requer redução de tamanho. - Alta variabilidade na composição.	20 - 50	0,3 – 0,6
Resíduos alimentares (FW)	- Produzido por hotéis, restaurantes, mercados e empresas de processamento de alimentos. - Requer redução de tamanho. - Alta variabilidade na composição. - Facilmente digerível, pode gerar inibição por acidificação.	5 - 30	0,3 – 0,8

Tabela 2 - Biomassas geralmente utilizadas em processos de digestão anaeróbia e suas características principais. Fonte: Adaptação Vasco Correa (2018). - Continuação

Tipo de substrato	Características	TS [%]	Produção de biogás [m³/kg VS]
Resíduos agrícolas e de culturas energéticas	<ul style="list-style-type: none"> - Abundante em disponibilidade. - Composto principalmente por celulose, hemicelulose e lignina e/ou amido. - Os resíduos incluem folhas, palha, canudos, etc. - As culturas energéticas podem ser ensiladas para armazenamento. - Altamente recalcitrante: necessita de pré-tratamento para melhorar a digestibilidade. 	20 - 80	0,2 – 0,5
Lodo de esgoto	<ul style="list-style-type: none"> - Subproduto do tratamento de efluentes. - Rico em sólidos, patógenos e nutrientes. - Altamente contaminante: precisa de estabilização. - Baixa digestibilidade: pré-tratamento ou codigestão com outras matérias-primas podem melhorá-la. 	20 - 35	0,8 – 1,2
Algas	<ul style="list-style-type: none"> - Macroalgas (algas marinhas) ou microalgas. - Baixa digestibilidade, estrutura recalcitrante, especialmente se a parede celular não foi rompida. Poderia melhorar com pré-tratamento. 	2 - 30	0,5 – 0,8

Portanto, a produção de biogás pode ser gerada a partir de muitos tipos de resíduos e essa composição varia conforme o substrato (matéria orgânica) utilizado e as condições do processo, tornando difícil a padronização (Galbiatti et al.,2010). Uma forma de aproveitar esses resíduos é através dos biodigestores, que são unidades de tratamento de resíduos orgânicos que utilizam a digestão anaeróbia para produzir biogás e biofertilizante. No entanto, para estimar a produção de biogás e de metano, é necessário realizar experimentos e métodos analíticos, que podem ser demorados e custosos. Por isso, também é fundamental ponderar a aplicação de modelos matemáticos para prever o desempenho dos sistemas de biodigestão, avaliar diferentes cenários e otimizar as condições operacionais (Guimarães et al.,

2018; Matos et al., 2017). Desta forma, é importante que haja o suporte de ferramentas computacionais que sejam capazes de otimizar o seu sistema. A construção de modelos matemáticos para processos de tratamentos biológicos apresenta muitas dificuldades no que diz respeito à complexa abordagem cinética e ao seu comportamento não linear. Recentemente, ferramentas de Inteligência Artificial (IA) foram estudadas como formas adequadas e precisas para resolver problemas de engenharia. Dentre essas ferramentas, a modelagem de Redes Neurais Artificiais (RNAs) é amplamente aceita devido à sua capacidade para resolver problemas complicados. Sua precisão depende da disponibilidade de conjuntos de dados de entrada e saída suficientes (Ghatak et al., 2018). Opções como realização de ensaios em laboratório mostram a necessidade e a importância de se desenvolver uma ferramenta computacional de apoio que traga agilidade e confiabilidade. Esta agilidade poderá servir para propor uma mudança em um processo de biodigestão que possa aumentar sua eficiência de produção de biogás e, conseqüentemente, na geração de energia elétrica caso esse biogás seja utilizado em um sistema de cogeração.

1.2 Objetivos

No presente estudo, a partir de dados experimentais coletados da literatura mundial sobre o processo de produção de biogás, de metano e a percentagem de metano na mistura, respectivamente, o objetivo principal é desenvolver modelos usando redes neurais artificiais (RNA) visando obter uma predição robusta em amplas faixas de operações desses parâmetros. Os parâmetros avaliados foram a natureza da biomassa, o tipo de reator com o biodigestor modo de alimentação, o teor de sólidos voláteis (VS), o pH, a taxa de carga orgânica (*Organic Loading Rate*, OLR), o tempo de retenção hidráulica (*Hydraulic Retention Time*, HRT), a temperatura e o volume do reator. Os modelos desenvolvidos buscaram prever: a produção de biogás, a produção de biometano e sua concentração na mistura. Para assegurar que as performances estejam adequadas, as RNA foram avaliadas pela sua capacidade de estimação através do coeficiente de determinação (R^2) e também através da soma do erro quadrático (SSE) obtidos.

Dessa forma, pode-se apresentar os seguintes objetivos específicos:

1. Construção de um banco de dados robusto para desenvolvimento de modelos baseados em dados científicos da literatura.
2. Análise do banco dos dados levantados na etapa anterior e construção de histogramas para remoção dos dados atípicos.
3. Desenvolvimento de modelos matemáticos para produção acumulada de biogás, produção acumulada de metano e concentração volumétrica de CH₄, respectivamente, por meio de redes neurais artificiais.
4. Avaliação de possíveis regiões com maior robustez.
5. Construção de gráficos de superfície de resposta para determinar os valores de resposta desejáveis e as condições ótimas de operação, bem como para identificar interações e curvaturas significativas entre as variáveis.

1.3 Estrutura do trabalho

Na seção 1, há uma introdução à temática abordada no presente trabalho, indicando a contextualização e objetivos a serem alcançados.

Na seção 2, será apresentada uma revisão bibliográfica, com o intuito de aprofundar os conceitos relacionados ao processo de produção de biogás e metano, incluindo os fatores que influenciam a produção de biogás, suas características e aplicações, além de fundamentos sobre as Redes Neurais Artificiais (RNA).

Na seção 3, será apresentada a metodologia adotada na elaboração dos bancos de dados e dos modelos de Redes Neurais Artificiais (RNA).

Na seção 4, serão apresentados os resultados e discussão do presente estudo, detalhando a divisão dos dados de treinamento, teste e validação, a escolha da melhor topologia da RNA, a RNA eleita e superfícies de resposta, a validação do modelo – incluindo análises e perspectivas para melhoria dos modelos – e, por fim, uma avaliação da robustez do modelo.

Por último, na seção 5, são indicadas as conclusões do presente trabalho, com perspectivas de trabalhos futuros.

2 Revisão bibliográfica

2.1 Processo de formação do biogás

A transformação da matéria orgânica em diversas substâncias químicas, no decurso da digestão anaeróbia, processa-se através de uma cadeia de degradações sucessivas devidas a diferentes tipos de microrganismos. A formação do biogás envolve basicamente três etapas, sendo elas a fermentação, dividido em duas fases, a hidrólise e a acidogênese, a acetogênese e a metanogênese (Van Haandel & Lettinga, 1994).

Na primeira fase, as bactérias hidrolíticas convertem a matéria orgânica em moléculas menores, ou seja, proteínas em peptídeos e aminoácidos, polissacarídeos em monossacarídeos, gorduras em ácidos graxos. Essa degradação é feita pela ação de enzimas extracelulares, uma vez que os polímeros orgânicos, por serem de tamanho consideravelmente grande, não conseguem penetrar no interior das células bacterianas. Em seguida, durante a acidogênese, as bactérias transformam os produtos da fase anterior em ácidos solúveis, álcoois e outros compostos (Chernicharo, 1997).

Durante a acetogênese, ocorre a conversão dos produtos obtidos na fermentação em ácidos acéticos (CH_3COOH), hidrogênio (H_2) e dióxido de carbono (CO_2). Esses compostos são utilizados como substratos pelas arqueas na fase subsequente do processo (Chernicharo, 1997). Na fase da metanogênese ocorre a formação de metano, decorrente da ação das bactérias metanogênicas, as quais são obrigatoriamente anaeróbias e extremamente sensíveis a mudanças na temperatura e no pH. As arqueas metanogênicas transformam o hidrogênio (H_2), o dióxido de carbono (CO_2) e o ácido acético (CH_3COOH), obtidos na etapa anterior, em metano (CH_4) e dióxido de carbono (CO_2) (Van Haandel & Lettinga, 1994), conforme verificado na Figura 2.

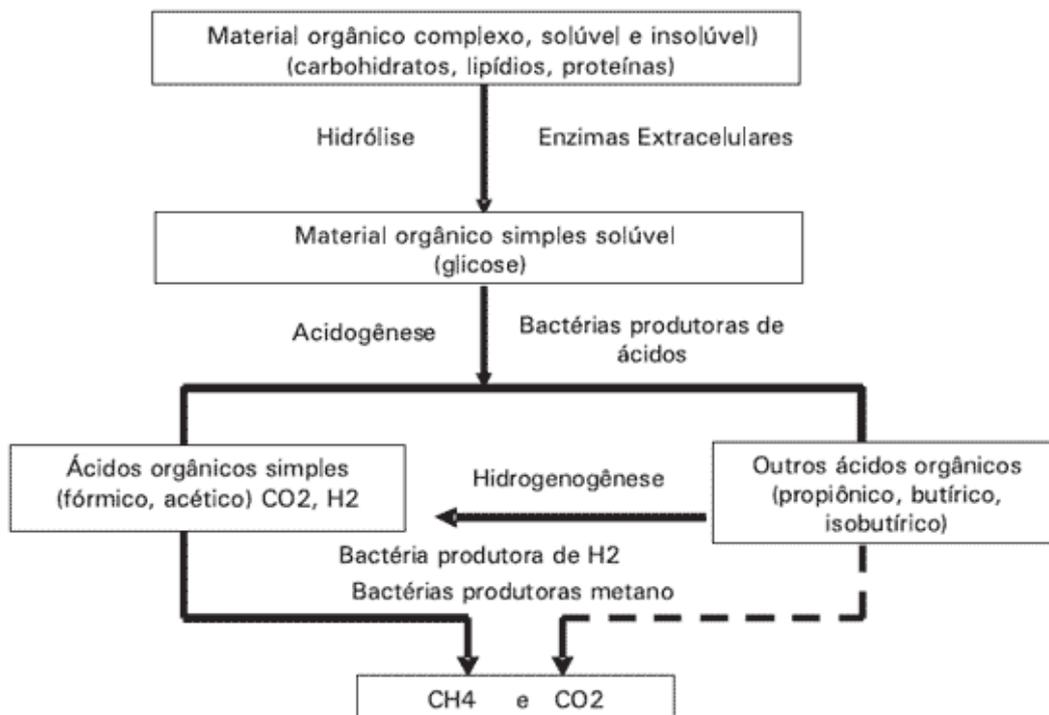


Figura 2 - Etapas metabólicas do processo de digestão anaeróbica (Adaptado de Oliveira e Ramalho, 2006).

Cada etapa da digestão anaeróbica possui condições operacionais específicas que favorecem o crescimento e a atividade dos micro-organismos envolvidos. Essas condições incluem fatores como temperatura, pH, umidade, tempo de retenção hidráulica, carga orgânica, relação carbono/nitrogênio, relação acetato/propionato, inibidores e nutrientes. A otimização dessas condições permite aumentar a eficiência e a estabilidade do processo, bem como a produção e a qualidade do biogás. Por outro lado, o desequilíbrio ou a variação dessas condições pode causar problemas como acidificação, inibição ou paralisação do processo. Os valores de pH e de temperatura durante as fases do processo de formação do biogás variam de acordo com o tipo de micro-organismo envolvido e o substrato utilizado. De modo geral, pode-se dizer que na fase de hidrólise, o pH ideal fica em torno de 5,2 a 6,3 e a temperatura pode ser mesofílica (entre 25°C e 40°C) ou termofílica (entre 50°C e 60°C), dependendo do substrato e do tipo de biodigestor (Portal do biogás, 2023). Na fase de acidogênese, o pH ideal fica entre 5,5 e 6,0 e a temperatura pode ser mesofílica ou termofílica, sendo que a faixa termofílica favorece a produção de ácidos graxos voláteis (Amaral; Steinmetz; Kunz, 2020). Na fase de acetogênese, o pH ideal fica entre 6,0 e 7,0 e a temperatura pode ser mesofílica ou

termofílica, sendo que a faixa termofílica favorece a produção de acetato (Amaral; Steinmetz; Kunz, 2020). Já na fase de metanogênese, o pH ideal fica entre 6,7 e 7,5 e a temperatura pode ser mesofílica ou termofílica, sendo que a faixa termofílica favorece a produção de metano (Portal do biogás, 2023; Amaral; Steinmetz; Kunz, 2020).

2.2 Fatores que influenciam a produção de biogás

Existem diversos fatores que influenciam a produção de biogás em um biodigestor. Os principais parâmetros e que são objeto de estudo desse trabalho são: tipos de substrato, temperatura, pH, tempo de retenção hidráulica (HRT), carga orgânica (OLR), sólidos voláteis (VS), volume do reator, geometria do reator e tipo de abastecimento. Demirel e Scherer (2008) revisaram o potencial de conversão anaeróbica de vários substratos, incluindo resíduos de alimentos e resíduos de frutas e vegetais, para a produção de biogás. Os autores analisaram aspectos cruciais relacionados à composição desses substratos, como teor de matéria orgânica, relação carbono/nitrogênio e presença de inibidores, os quais impactam diretamente a eficiência e a taxa de produção de biogás. Além disso, a revisão abordou os diferentes processos de digestão anaeróbica aplicados a esses resíduos, como a biodigestão em reatores de fluxo contínuo, reatores de batelada e biodigestores de lagoas. Os autores enfatizaram a importância de ajustar as condições operacionais, como temperatura, pH e tempo de retenção hidráulica, para maximizar a produção de biogás e otimizar a estabilidade do processo. Ji et al. (2017) revisaram a digestão anaeróbica de resíduos de frutas e vegetais para produção de biogás, discutindo as características dos resíduos e os fatores que afetam o desempenho do processo de digestão anaeróbica. Os autores concluem que a digestão anaeróbica de FVW é uma tecnologia promissora para o tratamento de resíduos orgânicos, a produção de energia limpa e a recuperação de nutrientes, mas que ainda requer mais pesquisas para otimizar o processo, melhorar a qualidade dos produtos e aumentar a viabilidade técnica e econômica. A seguir, serão apresentados os parâmetros de maior relevância para este estudo.

O tipo e a composição de substrato a ser utilizado afeta a estabilidade e a eficiência do processo e a qualidade do biogás, o que está diretamente ligado à quantidade de nutrientes e contaminantes potenciais (metais, patógenos, contaminantes orgânicos contidos na matéria orgânica). A escolha do material certo influencia no resultado do processo, na maximização da produção de energia e na boa qualidade de biogás (Karlsson, 2014).

O substrato deve conter vários componentes que são necessários para a atividade dos microrganismos nos processos enzimáticos, tais como oligoelementos e vitaminas. Na produção de biogás, um fator de extrema importância a ser levado em conta é a relação carbono: nitrogênio (C:N). É importante que essa relação não esteja muito baixa e que a proporção de nitrogênio seja inferior à de carbono. O processo pode ser facilmente afetado em razão da inibição das bactérias causada pela liberação de amônia, causando deficiência de nitrogênio no processo. A proporção ideal de carbono e nitrogênio dependerá de cada tipo de substrato, bem como das condições do processo (Karlsson, 2014).

De acordo com a United States Environmental Protection Agency (EPA), a maioria dos resíduos orgânicos, como restos de comida, folhas e galhos, estão na forma sólida. Em substratos para produção de biogás, os sólidos totais (TS) são a soma de todas as substâncias sólidas presentes no material, incluindo a fração orgânica e inorgânica. Sólidos voláteis (VS) é uma medida do conteúdo de matéria orgânica em um material. Ela é obtida pela perda de massa de um material após a sua combustão a uma temperatura elevada. Os sólidos voláteis incluem a fração do material que pode ser decomposta e convertida em gases durante o processo de digestão anaeróbia, e, portanto, é um parâmetro importante para o dimensionamento e operação de sistemas de tratamento de resíduos. (Karlsson, 2014).

De acordo com Zhang et al. (2014), os sólidos totais (TS) e sólidos voláteis (VS) de restos alimentares (*food waste*, FW) estavam nas faixas de 18,1–30,9 e 17,1–26,35, respectivamente, indicando que a água representa entre 70–80% em FW. Devido a este alto teor de umidade, o FW é um substrato orgânico facilmente biodegradável. Resíduos domésticos, por exemplo, nem sempre têm a mesma composição, o que varia conforme o lugar e a época do ano em que o resíduo é coletado. A composição dos resíduos varia dependendo de onde e em que condições são gerados (tipo de solo, clima etc.). Além disso, as condições de armazenamento

e potencial de aproveitamento para fins de reaproveitamento, reciclagem, dos resíduos também liberam gases na atmosfera (Karlsson, 2014). De acordo com Ji et al. (2017), os sólidos totais (TS) e sólidos voláteis (VS) de resíduos de frutas e vegetais (FVW) estavam nas faixas de 10,3–25,1 e 8,9–22,9, respectivamente.

Geralmente, a codigestão de diferentes materiais apresenta melhor resultado, pois gera maior quantidade de biogás do que se for adicionado ao biodigestor apenas um tipo de substrato. A explicação para isso é que, quando os substratos são misturados, tem maior chance de conter todos os componentes que são importantes para o crescimento microbiano. A mistura pode, por exemplo, proporcionar melhor relação C:N. Além disso, se o substrato é complexo promove, o crescimento de vários tipos de microrganismos no biodigestor (Karlsson, 2014).

O biodigestor pode ser classificado quanto ao modo de operação como contínuo ou descontínuo/batelada. O processo de digestão pode ser conduzido de várias formas, dependendo do substrato a ser digerido. Os biodigestores do tipo contínuo são alimentados continuamente ou em intervalos regulares com pequenas quantidades de substrato, mantendo um fluxo constante de entrada e saída do material. Esse tipo de biodigestor é mais complexo e requer um maior controle operacional, mas apresenta uma produção contínua e elevada de biogás (SANTOS et al., 2017). Tal processo é possível se o substrato possuir níveis de sólidos inferiores a 5%, como, por exemplo, águas residuais municipais e industriais. O conteúdo de sólidos no lodo é entre 5% e 15%, lamas provenientes de estações de tratamento de esgotos podem ser introduzidas continuamente no processo. Essa é a denominada digestão semicontínua, com cerca de 1 a 8 bombeamentos por dia. É normal dispor de teor de sólidos entre 20% a 25%, no início do processo, em materiais como resíduos vegetais e resíduos alimentares. Mediante a adição de materiais sólido-líquidos é possível que o material na entrada do biodigestor seja contínuo, o que é vantajoso para os microrganismos, pois haverá distribuição uniforme de material durante todo o dia. Isso facilita a interação entre os diferentes grupos de microrganismos na cadeia de degradação, e também reduz o risco de sobrecarga, caso se alimentem com uma grande quantidade de material em uma única vez. O fornecimento uniforme do substrato permite maior stress entre os microrganismos (Karlsson, 2014).

O oposto de um processo contínuo é um processo em batelada, ou seja, quando todo o material é alimentado de uma só vez e o material é sempre o mesmo durante todo o processo de degradação (Karlsson, 2014). A este sistema não é adicionado novo material e a produção de metano geralmente é mais elevada no início e diminui com o tempo. Quando o material é digerido, este é retirado do reator e uma nova quantidade de substrato é adicionada. Um exemplo de um processo descontínuo é um aterro, no qual o resíduo é tratado no mesmo lugar por um longo tempo. A digestão anaeróbia por meio do processo em batelada também é comum na produção de biogás em pequenas propriedades, principalmente nos países asiáticos. Esse sistema é vantajoso a partir de um material orgânico. Algumas vezes nesse processo pode ser difícil conseguir alta e constante produção de biogás, principalmente se o material tem alto teor de sólidos.

A configuração do processo é um fator importante para a eficiência da produção de metano a partir de resíduos biodegradáveis. Para isso, são utilizados processos de estágio único e dois estágios. Quando todas as fases da digestão anaeróbia ocorrem simultaneamente em um único reator, chamamos de reator de estágio único (Forster-Carneiro et al., 2008). No entanto, a estabilidade do digestor anaeróbio de estágio único para resíduos facilmente degradáveis é uma preocupação (Lee et al., 1999). Por outro lado, o digestor anaeróbio de dois estágios é utilizado para produzir hidrogênio e metano em dois reatores separados, sendo que a primeira fase é dedicada à produção de hidrogênio e ácido graxo volátil, enquanto a segunda fase é responsável pela conversão dos ácidos graxos voláteis em metano e dióxido de carbono (Chu et al., 2008). Foi relatado que a digestão anaeróbia em dois estágios proporciona uma operação mais eficiente em comparação com um único estágio. De acordo com Massanet-Nicolau et al. (2013), o rendimento de metano a partir de resíduos alimentares aumentou em 37% no processo de fermentação de metano em dois estágios. Os maiores rendimentos de metano a partir de resíduos alimentares foram relatados por Koike et al. (2009), que obtiveram uma produção de biogás de 850 L/g VS durante o processamento de produção de hidrogênio e metano em dois estágios de resíduos alimentares.

A temperatura é um parâmetro de grande importância que afeta a digestão anaeróbica, influenciando tanto a atividade enzimática quanto o rendimento de

metano e a qualidade do digestato (Zhang et al., 2014). Durante a degradação da matéria orgânica, o oxigênio liberado faz com que o composto aerado aqueça. No processo de produção de biogás, a quantidade de calor gerada é muito pequena em relação à energia liberada pela respiração celular, que está diretamente relacionada ao produto final, o metano. Para que os microrganismos se desenvolvam adequadamente e produzam quantidades satisfatórias de biogás, é necessário fornecer calor externamente (Karlsson, 2014).

Os microrganismos envolvidos na digestão anaeróbica são classificados em psicofílicos, mesofílicos e termofílicos, cada grupo com sua faixa de temperatura ideal. Para os microrganismos psicofílicos, a temperatura ideal é inferior a 25°C, o que resulta em uma decomposição mais lenta e baixa produção de biogás. Já os microrganismos mesofílicos apresentam um pico de crescimento em uma faixa de temperatura ideal entre 37°C e 42°C, resultando em rendimentos relativamente altos e boa estabilidade do processo, uma vez que a temperatura ideal para a formação de metano está entre 35°C e 37°C. Caso a temperatura de fermentação esteja abaixo do ideal, podem ocorrer produção de ácidos graxos e álcoois, prejudicando a formação de metano, diminuindo o pH e interrompendo o processo. Na digestão termofílica, os microrganismos termofílicos atuam em uma faixa de temperatura entre 50°C e 60°C, sendo que a temperatura de operação usual das plantas de biogás varia entre 50°C e 55°C. A atividade dos microrganismos na digestão termofílica é 25% a 50% mais elevada do que na digestão mesofílica (Karlsson, 2014).

A digestão anaeróbica apresenta um melhor desempenho em temperaturas mais elevadas, o que proporciona vantagens como taxas metabólicas mais altas, maior destruição de patógenos e maior produção de biogás (Zhang et al., 2014). A elevação da temperatura acelera a atividade metabólica dos microrganismos, desde que estejam adaptados às condições térmicas. Isso intensifica a degradação da matéria orgânica, resultando em uma redução mais rápida do volume do biodigestor (Karlsson, 2014). Estudos mostram que a digestão termofílica apresenta menos problemas de inibição devido ao acúmulo de amônia do que a digestão mesofílica (Gallert et al., 1997; Gallert et al., 1998). Além disso, a produção de biogás, a degradação do nitrogênio orgânico e a assimilação de fósforo aumentam com a temperatura (Wei et al., 2010; Sánchez et al., 2001). A viscosidade do meio

reacional também diminui em condições termofílicas, facilitando a agitação do sistema (Karlsson, 2014). Outra vantagem da digestão termofílica é a pasteurização natural do material, eliminando microrganismos causadores de doenças, como a Salmonella. O tempo de exposição do material a altas temperaturas determina a viabilidade desse processo. No entanto, a digestão termofílica pode ser suscetível a perturbações devido à proximidade da temperatura ótima dos microrganismos em relação à temperatura máxima suportada por eles, resultando em morte ou inibição (Karlsson, 2014).

Embora a digestão termofílica apresente várias vantagens, é importante considerar algumas desvantagens, como a sensibilidade a mudanças ambientais, especialmente quando a taxa de variação de temperatura excede $1^{\circ}\text{C}/\text{dia}$, o que requer que as mudanças de temperatura sejam inferiores a $0,6^{\circ}\text{C}/\text{dia}$ para manter a estabilidade da digestão (Zhang et al., 2014; Appels et al., 2011). Além disso, foi observado que a taxa global de solubilização de resíduos alimentares é significativamente menor em condições termofílicas do que em condições mesofílicas (Komemoto et al., 2009). Portanto, a temperatura desempenha um papel crucial na digestão anaeróbica, e a seleção adequada da faixa de temperatura ideal é fundamental para obter um processo eficiente e estável, com produção satisfatória de biogás e destruição de patógenos.

As variações de acidez seguem o mesmo princípio da temperatura, na qual cada grupo de bactérias possui seu ponto ideal de atividades enzimáticas. Nas fases da hidrólise e da acidogênese, o pH ideal fica em torno de 5,2 à 6,3. Já na acetogênese e metanogênese, o pH ideal fica entre 6,5 à 8,0 (FNR, 2010). Abaixo do pH crítico, pode haver precipitação dos íons metálicos, bem como inibição da ação bacteriana, devido à produção de ácidos, visto que a parede celular bacteriana é mais permeável às moléculas não dissociadas em comparação com suas formas ionizadas. Acima do pH crítico, a partir do qual a concentração de íons carbonato disponíveis é elevada, os metais pesados passam a ser precipitados na forma de carbonatos, e neste caso é acentuada a influência do pH (FNR, 2010).

A geração de biogás obtém maior rendimento em pH neutros ou ligeiramente superior (pH entre 7,0 e 8,5). A fim de manter um pH neutro e estável, é necessário que a alcalinidade do meio seja relativamente elevada e constante. A

alcalinidade é uma medida da quantidade de substâncias alcalinas (básicas) presentes no processo de geração de biogás. Quanto maior for a alcalinidade, maior será a capacidade tampão do processo, que, por sua vez, promove um pH estável (Karlsson, 2014).

O tempo de retenção hidráulica (*Hydraulic retention time*, HRT) ou é o tempo que o material permanece no interior do biodigestor, ou seja, tempo entre a entrada do efluente e a saída do afluente do digestor. O tempo de retenção hidráulica pode ser obtido pela relação entre o volume de mistura líquida do biodigestor e a carga diária. Durante o tempo de retenção, a matéria orgânica é degradada e ocorre a produção de biogás (Tshemese et al., 2023). De acordo com os diferentes substratos utilizados para a alimentação do biodigestor, e dos demais fatores, o tempo de retenção varia de 4 a 60 dias (Mazzucchi, 1980; Filho, 1981). Para o substrato composto de resíduos de matéria orgânica, o tempo de retenção situa na faixa de 20 a 30 dias, sendo que, com 30 dias, a carga orgânica já apresentou grande redução e, devido a isso, a produção de biogás já atingiu seu máximo e a partir desse ponto passa a decair (Beux, 2005).

O parâmetro relacionado a taxa de carga orgânica (*Organic Loading Rate* – OLR) é utilizado para medir a quantidade de substrato adicionada a um sistema de digestão constante por unidade de volume por dia, tornando-se relevante para reatores contínuos ou semi contínuos (Tshemese et al., 2023). Diversos estudos anteriores apontam que o OLR aplicado é um fator crucial na produção de biogás, que pode ser influenciado pela concentração do substrato e pelo tempo de retenção hidráulica. Entretanto, outras variações podem impactar o sistema de maneira diferente. A elevação da concentração de substrato, por exemplo, pode aumentar o OLR e melhorar a produção de biogás, mas altas concentrações podem levar a inibições no sistema ou alterações nas vias microbianas (Grangeiro et al., 2019).

Vários estudos têm sido realizados para determinar a taxa de carga ótima para diferentes tipos de resíduos e condições de operação. Em um experimento realizado por Nagao et al. (2012), que teve como objetivo determinar a taxa de carga orgânica máxima (OLR) para a digestão anaeróbia úmida de resíduos alimentares em um único estágio, a taxa volumétrica de produção de biogás aumentou à medida

que a OLR aumentava para 3,7, 5,5, 7,4 e 9,2 kg-VS m³/d, respectivamente, e foi mantida na mesma taxa. No entanto, na maior OLR (12,9 kg-VS m³/d), a taxa volumétrica de produção de gás diminuiu abaixo da taxa de produção de gás na OLR de 7,4 kg-VS m³/d. Em um estudo para comparar a biodigestão de restos alimentares autoclavados e não tratados (Tampio et al., 2014), o maior rendimento de metano foi obtido com taxa de carga orgânica de 3 kg de sólidos voláteis/m³/dia para restos alimentares não tratados e de 4 kg de sólidos voláteis/m³/dia para restos alimentares autoclavados. Agyeman e Tao (2014), que teve como objetivo avaliar os efeitos do tamanho das partículas de resíduos alimentares e da taxa de carga orgânica (OLR) na co-digestão anaeróbia de resíduos alimentares e esterco de gado leiteiro, relataram que a taxa de produção de biogás aumentou quando a OLR foi aumentada de 1 para 2 g de sólidos voláteis/L/dia e apenas ligeiramente quando a OLR foi aumentada de 2 para 3 g de sólidos voláteis/L/dia. O rendimento específico de metano atingiu o pico na OLR de 2 g de sólidos voláteis/L/dia nos digestores com restos alimentares finos e médios. Além disso, a co-digestão de restos alimentares com lodo ativado foi realizada em sistemas anaeróbios mesofílicos e termofílicos em diferentes OLR. O sistema termofílico apresentou a melhor capacidade de carga a uma OLR extremamente alta de 7 g de sólidos voláteis/L/dia, enquanto o sistema mesofílico apresentou a melhor estabilidade do processo em baixas OLRs (<5 g de sólidos voláteis/L/dia) (Carlsson et al., 2008). Portanto, é importante determinar a taxa de carga ótima para cada tipo de resíduo e condições de operação, a fim de maximizar a produção de biogás e evitar o colapso do reator devido ao acúmulo excessivo de ácidos graxos voláteis (Hagen et al., 2014).

A produção de biogás é influenciada por diferentes fatores, como os citados anteriormente. Porém existem outros fatores que não foram objeto de estudo desse trabalho, mas que tem importância no processo. Esses fatores serão descritos a seguir:

- Tipo de inóculo: O inóculo é um fator importante que afeta a produção de biogás e biometano por meio da digestão anaeróbia (DA) de resíduos orgânicos. O inóculo fornece uma população inicial de microrganismos que degradam os substratos e produzem biogás. O tipo, a qualidade e a quantidade de inóculo podem influenciar o rendimento, a composição e a

taxa de produção de biogás (Rajput; Sheikh, 2019). O inóculo ideal deve ter uma alta atividade metanogênica, uma boa adaptação ao substrato, uma diversidade microbiana adequada e uma estabilidade operacional (Khadka et al., 2022). A relação entre o substrato e o inóculo (S/I) é um fator importante que afeta a produção de biogás. Uma relação S/I muito baixa pode causar uma inibição do processo por excesso de microrganismos, enquanto uma relação S/I muito alta pode causar uma deficiência de nutrientes e um desequilíbrio do pH (ALI et al., 2023). Uma relação S/I ótima depende do tipo e da composição do substrato e do inóculo, bem como das condições operacionais do reator.

- Concentração de nutrientes nos substratos: Os nutrientes são essenciais para o crescimento e a atividade dos microrganismos que realizam a DA de resíduos orgânicos e produzem biogás e biometano. Os nutrientes mais importantes são o carbono, o nitrogênio, o fósforo, o potássio e os oligoelementos. A relação entre o carbono e o nitrogênio (C/N) é um fator crítico que afeta a produção de biogás, pois influencia o equilíbrio do pH, a disponibilidade de substrato e a inibição por amônia. Uma relação C/N ótima varia entre 20 e 30, dependendo do tipo e da composição dos resíduos alimentares. Uma relação C/N muito alta pode causar uma acidificação excessiva do meio, enquanto uma relação C/N muito baixa pode causar uma inibição por amônia (Khadka et al., 2022). O fósforo e o potássio são necessários para a síntese de biomassa e a manutenção do metabolismo celular. A relação entre o fósforo e o nitrogênio (P/N) também é importante para a produção de biogás, pois influencia a atividade enzimática, a transferência de elétrons e a estabilidade do processo. Uma relação P/N ótima varia entre 0,03 e 0,06, dependendo do tipo e da composição dos resíduos alimentares. Uma relação P/N muito alta pode causar uma precipitação de fosfato de amônio, enquanto uma relação P/N muito baixa pode causar uma deficiência de fósforo (Ali et al., 2023). Os oligoelementos são necessários em pequenas quantidades para atuar como cofatores de enzimas ou como componentes estruturais de proteínas. Os oligoelementos mais importantes para a produção de biogás são o ferro, o níquel, o cobalto, o molibdênio, o zinco, o selênio e o tungstênio. A deficiência ou o excesso

de oligoelementos pode causar uma redução ou uma inibição da produção de biogás, pois afeta a atividade metanogênica, a formação de cofatores ou a competição por sítios ativos (Rajput; Sheikh, 2019). Portanto, os nutrientes têm uma influência significativa na produção de biogás e biometano a partir de resíduos alimentares, e devem ser monitorados e controlados para otimizar o processo de DA.

- **Ácidos Graxos Voláteis (AGVs):** Segundo Paritosh et al. (2017) a produção e acúmulo de ácidos graxos voláteis (AGVs) pode mostrar efeitos inibitórios e prejudiciais sobre o processo de digestão anaeróbia, que pode levar a uma produção lenta de biogás. A inibição dos AGVs sobre a atividade dos metanógenos é causada por uma queda do pH, que pode levar à perda de atividade de enzimas sensíveis ao ácido. Além disso, altos níveis de ácidos não dissociados, que podem penetrar nas membranas celulares, podem danificar macromoléculas. Nas condições ótimas requeridas para atividade metabólica, os AGVs variam entre 2000–3000 mg/L (Paritosh et al., 2017).

2.3 Características e purificação do biogás

Os principais constituintes do biogás são o metano e o dióxido de carbono, havendo também baixas concentrações de monóxido de carbono, sulfeto de hidrogênio, amônia, entre outros (CASSINI, 2003). A Tabela 2 apresenta a porcentagem de cada constituinte presente no biogás.

Tabela 3 - Possíveis percentuais dos constituintes do biogás. Fonte: Adaptado de Cassini (2003).

Metano (CH₄)	50% a 75%
Dióxido de carbono (CO₂)	25% a 40%
Hidrogênio (H₂)	1% a 3%
Nitrogênio (N₂)	0,5% a 2,5%
Oxigênio (O₂)	0,1% a 1%
Gás Sulfídrico (H₂S)	0,1% a 0,5%
Amônio (NH₃)	0,1% a 0,5%
Monóxido de Carbono (CO)	0% a 0,1%

A quantidade de energia que o biogás pode fornecer depende principalmente da quantidade de metano (CH₄) que ele contém, pois o metano é o gás mais

combustível da mistura. A Tabela 3 mostra as características de alguns gases que podem estar presentes no biogás, como o dióxido de carbono (CO_2), o nitrogênio (N_2), o oxigênio (O_2), o monóxido de carbono (CO), o hidrogênio (H_2) e o sulfeto de hidrogênio (H_2S). Alguns desses gases, como o CO , o H_2 e o H_2S , também podem queimar e liberar energia, mas em menor quantidade do que o metano. Outros gases, como o CO_2 , o N_2 e o O_2 , não queimam e não contribuem para a energia do biogás. Além disso, alguns gases podem ter efeitos negativos, por serem tóxicos, asfixiantes ou corrosivos.

Tabela 4 - Características dos gases presentes no biogás. Fonte: Adaptado de Gandolla et al., 1997.

Espécie	Densidade (kg/m³)	Potencial calorífico (kJ/m³)	Limite explosivo	Solubilidade água (g/l)	Propriedades gerais
CH₄	0,717	35,600	5-15%	0,0645	inodoro, incolor, não tóxico, asfixiante e inflamável
CO₂	1,977	---	---	1,688	inodoro, incolor, asfixiante
N₂	1,250	---		0,019	inodoro e incolor
O₂	1,429	--		0,043	inodoro e incolor
CO	1,250	12,640	12,5-74%	0,028	Inodora, incolor, tóxico, inflamável
H₂	0,090	10,760	4-74%	0,001	inodoro, incolor, não tóxico, inflamável
H₂S	1,539	--	4,3- 45,5%	3,846	inodoro, muito tóxico

As impurezas do biogás o tornam muito corrosivo, principalmente devido a presença de gás sulfídrico, que entre outros materiais, ataca cobre, latão e aço (CRIBARI, 2004). Existem diferentes alternativas de purificação do biogás, devendo ser definida a mais adequada à aplicação energética desejável. A utilização do biogás em unidades de cogeração é o caso típico onde deve haver a remoção dos contaminantes como H_2S , hidrocarbonetos halogenados e siloxinas. A maioria dos fabricantes de turbina a gás especifica limites máximos para essas substâncias (VELÁZQUEZ et al., 2003). Segundo ZANETTE (2009), para a aplicação do biogás tanto no ramo automotivo como na rede de gás natural, deve-se remover

também o CO₂. Existem diversas tecnologias disponíveis para a purificação do biogás com vistas a atender as especificações de combustível veicular ou do gás natural. Na remoção de CO₂, pequenas quantidades de metano são removidas, acarretando em perdas, as quais devem ser minimizadas por razões econômicas e ambientais. A Tabela 4 apresenta um resumo das principais substâncias que devem ser removidas do biogás de acordo com sua utilização.

Tabela 5 - Necessidade de remoção de compostos gasosos de acordo com a utilização do biogás.

Fonte: Adaptado de Littig (2011).

Aplicação	H₂S	CO₂	H₂O
Microturbinas	Sim	Não	Sim
Motores Combustão Interna	Sim	Não	Sim
Caldeiras	<1000 ppm	Não	Não
Fogões	Sim	Não	Não
Combustível veicular	Sim	Recomendável	Sim
Rede de gás natural	Sim	Sim	Sim

Existem diversos mecanismos básicos que estão envolvidos para proporcionar a separação seletiva de componentes gasosos, entre eles estão a absorção física ou química, adsorção em uma superfície sólida, separação em membranas, separação criogênica e remoção de H₂S pelo óxido de ferro. Para o tratamento do biogás, os métodos de absorção física e química são geralmente utilizados por serem efetivos mesmo para as baixas vazões encontradas normalmente nas plantas de biogás. Além disso, o método é menos complicado, requer uma infraestrutura mais simples e é custo-efetivo (Zanette, 2009; Silva 2009).

2.4 Mercado do biogás no Brasil

O biogás começa a ganhar impulso no Brasil a partir dos choques no preço do petróleo promovidos pela Organização dos Países Exportadores de Petróleo (Opep), na década de 1970. Como o país, à época, importava cerca de 80% de seu consumo interno, buscou-se incrementar o uso de energias alternativas. O biogás era considerado uma alternativa viável para a diminuição de derivados de petróleo na matriz energética (Karlsson et al., 2014). Assim, foi criado em 1977 o Projeto

de Difusão do Biogás, no âmbito da Empresa Brasileira de Tecnologia e Extensão Rural (Emater), executado no estado de São Paulo e no Distrito Federal. A ausência de mão de obra especializada dificultou o atingimento dos objetivos do programa (Karlsson et al., 2014).

Em 1982, o Governo Federal passou a estimular a construção de biodigestores, no bojo do Programa de Mobilização Energética (PME) – Decreto 87.079, de 2 de abril de 1982 – usando materiais simples e de baixo custo. Embora a Emater do Paraná tivesse estimado que havia cerca de três mil biodigestores em operação no Brasil até 1984, foram inúmeras as dificuldades encontradas para difundir o uso do biogás. Faltava conhecimento técnico sobre a construção e operação dos biodigestores. A ausência de equipamentos devidamente projetados levava à adaptação de equipamentos, acarretando baixa durabilidade, baixa eficácia e custo elevado. Essas dificuldades levaram à perda de confiança nos sistemas de produção de biogás e ao progressivo abandono do uso de biodigestores (Cibiogás, 2020).

O biogás voltou ao foco nos anos 1990, dessa vez impulsionado pela questão ambiental, que passou a ocupar o centro das discussões sobre o aquecimento global e as mudanças climáticas decorrentes das emissões de gases de efeito estufa (GEE). A queima do biogás foi vista como uma forma de reduzir os GEE, por meio de projetos inseridos no chamado Mecanismo de Desenvolvimento Limpo (MDL), que possibilitava a venda de créditos de carbono gerados no processo. Vale ressaltar que o biogás não necessariamente era aproveitado como fonte de energia, bastando ser queimado em um flare. Nesse mercado, os produtores rurais não estavam integrados ao processo, sendo apenas fornecedores da biomassa em troca de uma percentagem da venda dos créditos. Como o processo não se sustentava economicamente, mais uma vez a biodigestão foi gradualmente abandonada (Mariane, 2018).

A despeito de suas vantagens enquanto fonte renovável de energia, o histórico de reveses do biogás no Brasil criou uma dificuldade de aceitação que só foi superada recentemente, em virtude do aprimoramento tecnológico e da ampliação da capacitação técnica do setor. De fato, inicialmente, utilizavam-se biodigestores convencionais que operavam com cargas orgânicas reduzidas e não retinham grandes quantidades de biomassa. Além disso, em geral, eram construídos

com materiais de baixo custo, o que ocasionava diversos problemas de operação e manutenção. Com o desenvolvimento tecnológico, não apenas esses equipamentos rudimentares foram sendo descartados e substituídos por equipamentos mais aprimorados, como também foram sendo desenvolvidos biodigestores de alta taxa, isto é, equipamentos de alto rendimento caracterizados por dispositivos que retêm grande quantidade de biomassa (Santos, 2019). O desenvolvimento de equipamentos específicos, de alta performance, desenhados para atender a situações concretas, em muito se beneficiou de experiências no exterior, notadamente na Alemanha e na Califórnia (Estados Unidos da América). O conhecimento técnico e a capacitação operacional foram se ampliando com base em incentivos que os governos locais concediam para a produção de biogás (Santos, 2019).

Um novo estímulo à expansão do biogás no Brasil veio com a implementação do Programa de Incentivo às Fontes Alternativas de Energia Elétrica (Proinfa) – Lei 10.438, de 26 de abril de 2002, e Lei 10.762, de 11 de novembro de 2003 – que objetivava ampliar a geração de energia elétrica por meio de pequenas centrais hidrelétricas (PCH) e das fontes eólica e biomassa, em conformidade com a redução nas emissões de CO₂ previstas no Protocolo de Kyoto. No que se refere especificamente à biomassa, a tarifa fixada pelo Proinfa ficou aquém do que os investidores esperavam, o que resultou, no primeiro leilão, em uma contratação de apenas cerca de 60% dos megawatts reservados ao setor sucroalcooleiro.

A instituição da geração distribuída, pela Agência Nacional de Energia Elétrica (ANEEL), foi um impulso adicional à biodigestão, pois o excedente de energia elétrica gerado nas propriedades rurais passou a poder ser inserido nas redes de distribuição (Milanez et al., 2021). Além disso, o ambiente de negócios para investimentos no setor de biogás vem se tornando cada vez mais favorável. A Lei nº 14.134/2021 (Nova Lei do Gás) e o seu regulamento, Decreto nº 10.712/2021, estabelecem que qualquer gás que não se enquadrar na definição de gás natural estabelecida na lei tenha tratamento equivalente, desde que aderente às especificações estabelecidas pela ANP — o que é o caso do biometano — criando assim segurança jurídica para os negócios que podem se utilizar da base legal e regulatória do gás natural (EPBR, 2022).

Em março, a Presidência da República publicou o Decreto nº 11.003/2022, que institui a Estratégia Federal de Incentivo ao Uso Sustentável de Biogás e Biometano. A Estratégia tem os seguintes objetivos: incentivar programas e ações para reduzir as emissões de metano, fomentar o uso de biogás e biometano como fontes renováveis de energia e combustível e contribuir para o cumprimento dos compromissos assumidos pelo país em convenções internacionais relacionadas ao tema (EPBR, 2022). Na esteira do Decreto 11.003, foram publicadas a Portaria do Ministério do Ambiente (MMA) 71, que instituiu o Programa Metano Zero (Programa Nacional de Redução de Emissões de Metano), e a portaria Normativa do Ministério de Minas e Energia (MME) 37, que inclui projetos de produção de biometano no rol do Reidi (Regime Especial de Incentivos para o Desenvolvimento da Infraestrutura) (iNFRA, 2022).

Dentre as medidas a serem desenvolvidas pelo Programa Metano Zero, se destacam: linhas de crédito e financiamento específicas para, entre outras ações:

- implantação de biodigestores;
- implantação de sistema de purificação de biogás, produção e compressão de biometano;
- criação de pontos e corredores verdes para abastecimento de veículos pesados movidos a biometano, tais como ônibus, caminhões e máquinas agrícolas;
- alavancagem da utilização ou desenvolvimento da tecnologia veicular;
- desoneração tributária para infraestruturas relacionadas com projetos de biogás e biometano.

Como resultado desses diversos fatores, a instalação de plantas de biogás, de todos os portes, ganhou um grande impulso nos últimos anos, como pode ser visto na Figura 3.

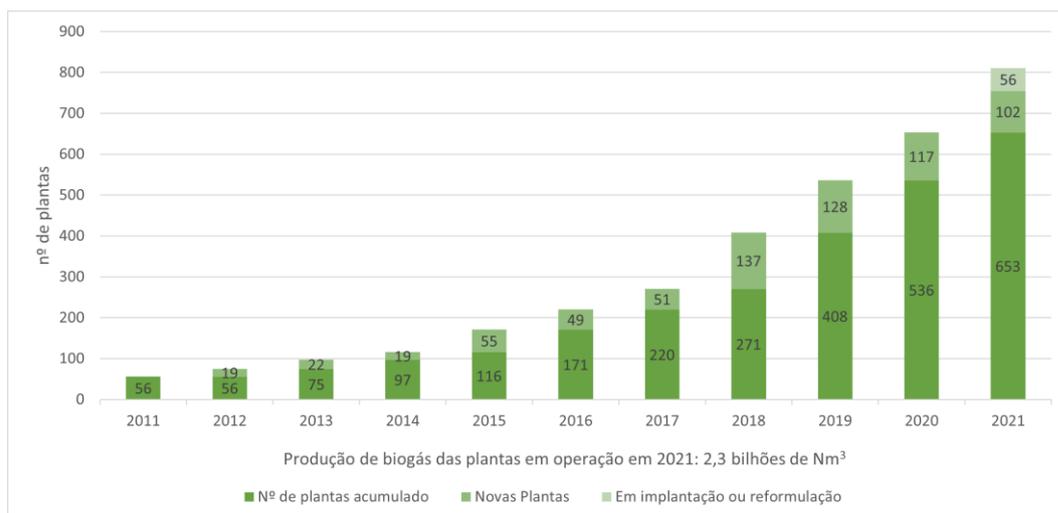


Figura 3 - Crescimento do número de novas plantas de biogás no Brasil na última década. Fonte: Elaboração própria. Adaptado de CIBiogás 2021

Segundo dados da ABiogás, o potencial teórico de produção de biogás brasileiro é de 84,6 bilhões de metros cúbicos por ano, o que seria suficiente para suprir 40% da demanda interna de energia elétrica e 70% do consumo de diesel (ABIogás, 2021). Considerando o cenário de 2,3 bilhões de metros cúbicos de biogás gerados em 2021 pelas 755 plantas em operação, o Brasil explora apenas 3% desse potencial. Para 2022, é esperado um crescimento de mais de 22% com a entrada em operação das 56 plantas que estão em fase de implantação ou em reforma. Com isso, espera-se alcançar uma produção anual de pelo menos 2,8 bilhões de Nm³ de biogás (CIBiogás, 2021).

De acordo com o último levantamento nacional de plantas de biogás e biometano, realizado pelo CIBiogás o número de unidades em operação passou de 653, em 2020, para 755 em 2021. Isso representa um crescimento de 16% a nível nacional. Os dados levantados no período contabilizam no total 811 plantas de biogás sendo que destas, 755 encontram-se em operação com aproveitamento energético (93%), 44 encontram-se em fase de implantação (5%) e 12 estão passando por reformulação ou reforma (2%).

Os substratos utilizados para produção de biogás no Brasil estão divididos em 3 categorias quanto sua fonte. A agropecuária envolve atividades de criação animal (aves, bovinos, caprinos, suínos, entre outros) e tem como substrato esterco animal, efluente proveniente do manejo dos dejetos (urina, fezes, água de lavagem etc.), restos de ração, carcaça de animais mortos não abatidos, entre outros. A

Indústria (agroindústrias) contempla abatedouros e frigoríficos, usinas de açúcar e etanol, fecularias e amidonarias, cervejarias, indústrias de óleo, vegetal, gelatina, entre outras e como substrato tem o efluente industrial e demais resíduos orgânicos provenientes do processo industrial. Já o saneamento contempla os aterros sanitários (RSU), as usinas de tratamento de resíduos orgânicos e as estações de tratamento de esgoto (ETE). Os substratos contemplam resíduos sólidos urbanos (RSU) depositado em aterro sanitário, RSU segregado na fonte (fração orgânica), resíduos de frutas e Vegetais (CEASA), restos alimentares (restaurantes, supermercado), esgoto sanitário (CIBiogás, 2021). É importante destacar que os resíduos alimentares (FW) e os resíduos de frutas e vegetais (FVW) são componentes importantes do lixo urbano, que representam uma fonte potencial de energia renovável e de fertilizantes orgânicos. No entanto, o descarte inadequado desses resíduos pode causar problemas ambientais e sanitários, como mau cheiro, contaminação da água, do solo e do ar, e riscos à saúde humana e animal. Dessa forma, a utilização desses substratos contribui para o saneamento ambiental, para a mitigação das mudanças climáticas, para a segurança alimentar e para a economia circular.

Quando se observam as fontes de substratos utilizadas pelos sistemas de biodigestão no Brasil, fica clara a prevalência de substratos de origem agropecuária na produção de biogás, que tem 80% no número de plantas. Contudo, embora os substratos agropecuários sejam majoritários, as plantas de biogás que os utilizam são responsáveis por apenas 10% do volume total de biogás produzido. A maior parte de biogás gerado, cerca de 74% do volume total, tem origem nas plantas que processam resíduos sólidos urbanos ou efluentes das estações de tratamento de esgoto (saneamento), sendo esses responsáveis por 9% no número de plantas. Os 16% restantes são produzidos por resíduos da indústria, que contemplam 11% do número total de plantas de biogás. (CIBiogás, 2021).

Já quando é observado o potencial de produção a curto prazo, ou seja, levando em conta apenas resíduos e efluentes que não apresentam obstáculos ao acesso imediato dentro das cadeias da pecuária, da indústria e do saneamento – estimado em 10,8 bilhões de metros cúbicos por ano (Instituto I17, 2021) – constata-se que 79% desse potencial ainda pode ser explorado. Há diversos estudos que tentaram quantificar esse potencial. Alguns buscaram definir potenciais

teóricos, construídos por meio da definição de parâmetros de cálculo. Outros tentaram, com base em experimentos empíricos, extrair os volumes produzidos de substratos distintos. Em comum, está a tentativa de se quantificar as possibilidades de produção do biogás utilizando biomassas diversas (Milanez et al., 2021).

2.5 Fundamentos sobre redes neurais artificiais (RNAs)

Segundo Park e Lek (2016), as redes neurais artificiais (RNAs), ou redes neurais, são uma técnica de modelo computacional complexa que pode ser usada sob aprendizado de máquina para prever a saída de um processo. A estrutura da rede neural artificial é semelhante às redes neurais biológicas em nossos cérebros humanos na forma como o sinal dos neurônios é transferido de um nó para outro. Os neurônios são os elementos de processamento em uma RNA, que recebem os dados de entrada como sinais, processam e produzem um sinal de saída. A Figura 4 mostra uma analogia entre a estrutura da RNA e uma célula neural. Os pesos (w_1 , w_2 , w_3) multiplicam os dados de entrada (x_1 , x_2 , x_3) e somam com um *bias* (b) antes de passar por uma função de ativação. Depois, essa resposta é conectada ao próximo neurônio, onde pode ser a saída final ou a entrada para outro neurônio.

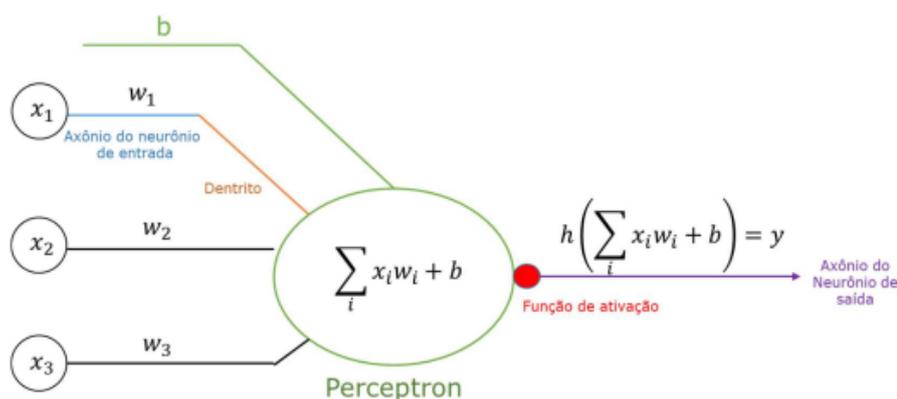


Figura 4 - Modelo artificial do neurônio. Adaptado de Araújo, 2020

O *bias* (b) é uma variável que ajusta a influência da entrada na ativação do neurônio. Ele permite que o neurônio tenha uma saída não nula, mesmo quando a entrada é nula. A função de ativação, por sua vez, define a amplitude do sinal que

deve avançar na rede para gerar a saída. O *bias* é somado na função ativação e atualizado como os pesos sinápticos (Ludwig JR. e Costa, 2007).

A rede é formada por neurônios que se conectam em camadas que formam um tipo de grafo. A arquitetura de uma rede tem as seguintes camadas: camada de entrada onde os dados de entrada são informados, camadas ocultas onde as conexões têm pesos e a maior parte do processamento é feita, essa camada tem o objetivo de extrair as informações relacionadas ao sistema e a camada de saída onde o resultado final é apresentado. Em uma arquitetura de camada única como da Figura 5, a conexão entre os nós pode ser mostrada na Equação 1 (Khalil et al., 2005):

$$Y_i = f \left(\sum_{j=1}^N W_{ij} X_j + b_i \right) \quad (1)$$

Na Equação 1 Y é o valor de saída, e X é a camada de entrada e W e b representam os pesos e o *bias*, respectivamente. O subscrito i e j denotam a camada anterior e a camada atual, respectivamente. A função, f é a função de ativação ou função de transferência da arquitetura de camada única. Existem vários tipos de funções de ativação de entrada e saída que podem ser usadas para modelar um único neurônio, as três mais comumente usadas são a função de transferência tangente hiperbólica (*tansig*), sigmóide logística (*logsig*) e função de transferência linear (*purelin*) (Neto, 2021). A escolha do tipo de função depende do tipo de projeto.

2.5.1 Arquiteturas das RNA's

Uma RNA consiste em uma conexão de neurônios que podem estar organizados em camadas através de sinapses. A estrutura e a funcionalidade de uma RNA dependem da escolha do número de camadas, dos pesos sinápticos e das funções de ativação. Silva et al. (2010) definem a arquitetura de uma RNA como “a maneira como vários neurônios estão dispostos, ou relacionados, uns com os outros”. As RNA's permitem que se ajuste a quantidade de neurônios, camadas e outros parâmetros de acordo com a necessidade de cada problema. Conforme mencionado anteriormente, as redes neurais podem ser divididas em três partes principais: camada de entrada, camadas ocultas e camada de saída.

A finalidade das RNAs é ajustar pesos e *biases* a partir dos dados de entrada e gerar um dado de saída igual ou próximo ao que foi usado para treinamento. Assim, quando uma nova amostra for fornecida à RNA treinada, ela produza uma saída adequada. Existem diversos tipos de redes, entre elas estão as redes neurais *feedforward*. Essas redes se caracterizam pela ausência de conexões de retroalimentação, ou seja, a saída dos neurônios de uma camada é usada como entrada para as outras camadas, seguindo uma ordem até a saída. Essa rede é capaz de modelar problemas em que os dados de entrada não têm dependência temporal com os dados de saída (Araujo, 2020). A rede mais usual do tipo feedforward é a de múltiplas camadas (*Multilayer Perceptron-MLP*). Em redes neurais multicamadas, os neurônios são dispostos em camadas, sendo que as camadas de entrada e saída são separadas por um conjunto de camadas intermediárias ou ocultas (*hidden layers*) (Aggarwal, 2018). Os modelos neurais artificiais mais populares e conhecidos são as arquiteturas do tipo MLP e será o foco do presente estudo. Seu funcionamento se inicia na camada de entrada, prossegue para as camadas ocultas e termina na camada de saída, fazendo com que os dados de entrada sejam sempre transmitidos para frente, de camada para camada. (Ludwig jr. e Costa, 2007).

A configuração das camadas e do número de neurônios em um sistema MLP depende de vários fatores, como por exemplo, dos valores iniciais fornecidos, da complexidade do problema, dos pesos sinápticos atribuídos inicialmente, entre outros. Segundo Cybenko (1997) o número de camadas ocultas necessárias pode ser determinado como: Uma camada oculta é suficiente para aproximação de funções contínuas e duas camadas ocultas são suficientes para aproximar qualquer função matemática. Em alguns casos, duas ou mais camadas podem facilitar no treinamento da rede, porém, a única camada que possui o valor exato do erro é a camada de saída, a camada anterior recebe apenas uma estimativa do valor, a anterior a esta recebe a estimativa da estimativa e assim por diante. Logo, quando o número de camadas aumenta demasiadamente, pode-se perder muito a precisão do erro e o mesmo se torna menos útil. (Braga et al., 2000). Ainda segundo Braga et al. (2000), o número de neurônios em cada camada oculta pode ser determinado de duas formas: baseando-se no número de entradas e saídas, ou considerando a quantidade de neurônios dez vezes menor que o número de padrões, sendo que cada padrão possui um conjunto de entradas e saídas. No entanto, outras alternativas

podem ser consideradas, pois essa escolha depende de vários fatores como já mencionado.

O número de neurônios nas camadas intermediárias de uma rede neural artificial afeta o desempenho do treinamento e a capacidade de generalização da rede. Dois problemas comuns relacionados à escolha do número de neurônios são o *underfitting* e o *overfitting*, que são situações de subajuste e sobreajuste da rede aos dados. O *underfitting* ocorre quando há poucos neurônios que não são suficientes para captar as características relevantes de um conjunto de dados complexo, resultando em um alto erro de treinamento. O *overfitting* ocorre quando há muitos neurônios que se ajustam demais aos dados de treinamento, memorizando os ruídos e as variações irrelevantes, e perdendo a capacidade de generalizar para novos dados. Portanto, é necessário encontrar um equilíbrio entre o número de neurônios e a complexidade dos dados (Sheela, 2013).

O número de neurônios nas camadas de entrada e saída é determinado pelo número de entradas e saídas, respectivamente, que a rede irá processar. Por exemplo, se a rede receber n valores de entrada, ela deverá ter n neurônios na camada de entrada, o mesmo vale para a camada de saída (Silva et al., 2010).

Essas funções possuem a responsabilidade de processar as entradas de cada neurônio, gerando uma saída, de acordo com os bias e pesos correspondentes do neurônio. A principal razão pelas quais as funções de ativação são utilizadas é para conferir uma capacidade não-linear ao processamento realizado pelas redes neurais. Isso é especialmente importante nas camadas ocultas da rede neural. Nas camadas de saída, as funções de ativação podem ter finalidades específicas, dependendo do problema que a rede neural está tentando resolver (Ceccon, 2020).

- Purelin: A função linear apenas aplica um fator de multiplicação ao valor que recebe. Apesar de cumprir todos os requisitos, essa função é limitada em sua capacidade de compreender relações mais complexas entre os dados. Na camada de saída, a função de ativação linear pode ser utilizada para tarefas simples, como em problemas de regressão, já que produz resultados em todo o domínio dos números reais (Ceccon, 2020). Pode-se definir a função como na Equação 2:

$$f(x) = x \quad (2)$$

- **Logsig:** A função logística ou sigmoide produz valores no intervalo $[0, 1]$. Ela é eficiente na questão da não-linearidade, e sua maior vantagem reside no fato de que o valor da derivada é máximo quando x está próximo de 0, o que tende a “empurrar” seu resultado para as extremidades do intervalo $[0, 1]$ ao longo do treinamento, o que é uma característica desejável, por exemplo, em problemas de classificação. Além disso, como se pode observar no gráfico da Figura 8, a função sigmoide não é centrada em zero. Na camada de saída, a função sigmoide é útil para produzir probabilidades em problemas de classificação binária, já que seus resultados, na faixa de $[0, 1]$, podem ser interpretados como a probabilidade de determinada instância pertencer ou não a determinada classe. A Equação 3 mostra como a função é definida:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

- **Tansig:** Similar a função sigmoide, a função tangente hiperbólica também tem um formato de ‘S’, mas varia de -1 a 1, em vez de 0 a 1 como na sigmoide. Ela se aproxima mais da identidade, sendo assim uma alternativa mais atraente do que a sigmoide para servir de ativação às camadas ocultas das RNAs (Facure, 2017). A definição da função está na Equação 4 a seguir:

$$f(x) = \frac{2}{(1 + e^{-2x})} - 1 \quad (4)$$

2.5.2 Treinamento das RNA's

O aprendizado de uma rede neural envolve algoritmos de otimização, para calibrar os parâmetros de pesos e *biases* da rede buscando a minimização de uma dada função objetivo. O treinamento de uma rede MLP é geralmente realizado pelo algoritmo *backpropagation*. Esse algoritmo, por meio do aprendizado supervisionado, ajusta os pesos sinápticos e *biases* da rede de acordo com o erro entre a saída desejada e a saída obtida. (Braga et al., 2000). O algoritmo contém duas fases: fase *forward* e fase *backward*.

No início da fase *forward*, o conjunto de dados de entrada é introduzido no modelo da rede na camada de entrada. Os dados de entrada juntamente com os pesos

passam para a camada oculta. A saída ponderada total é adicionada a um *bias* para produzir dados de entrada de um neurônio na camada de saída. Com base nos pesos, *bias* e função de transferência de cada neurônio, os sinais se propagam ao longo da rede, da camada de entrada para a camada de saída através da camada oculta, para gerar uma saída. Assim que o sinal de dados chega à saída da rede, o valor dos dados é comparado com o valor real dos dados experimentais de saída. A diferença ou erro entre o valor real e o valor previsto é calculado e os sinais de dados de erro são direcionados para trás, da camada de saída para a camada de entrada através da camada oculta. No decorrer da *backpropagation* dos sinais de dados, os pesos e o *bias* em cada neurônio das camadas são reformados para minimizar o sinal de erro dos dados na saída. Este procedimento iterativo é chamado de treinamento (Ghatak, 2018).

Um único passo de todos os vetores de treinamento usados para atualizar os pesos em uma rede neural é chamado de *epoch*. Durante o processo, a rede é alimentada com um dado de entrada e atualiza o peso da rede. Vários números de *epochs* são necessários para minimizar o erro entre os dados reais e os previstos. O número total de *epochs* necessárias para alcançar o grau de precisão é introduzido como a taxa de convergência. O erro pode ser minimizado através do ajuste de peso entre as camadas usando um algoritmo de otimização adequado (Ghatak, 2018). Vários algoritmos de treinamento de *backpropagation* têm sido usados para o treinamento de vários problemas de aproximação funcional e reconhecimento de padrões. Diversos trabalhos (Akbas et al., 2015; Yetilmezsoy et al., 2013; Behera et al., 2014; Jaroenpoj et al., 2014; Barik et al., 2015 e Palaniswamy et al., 2016) observaram que o algoritmo de treinamento de *Levenberg-Marquardt* é muito eficiente para a modelagem da RNA.

Após a conclusão efetiva do treinamento usando o conjunto de dados de treinamento, a arquitetura da rede neural pode ser armazenada e a rede é capaz de calcular novos valores de saída associados a um novo conjunto de dados de entrada. A arquitetura da rede neural armazenada precisa ser confirmada fornecendo um conjunto de dados de teste. Se o erro de teste for muito maior do que o erro de treinamento, a rede é observada como sobre ajustada para os dados. Uma rede neural ajustada com precisão pode dar erros aproximados de treinamento e teste. Uma vez que o treinamento e o teste da rede estejam concluídos, a rede é validada

com um novo conjunto de dados não utilizado (validação). Após obter confiança adequada no modelo, com base na validação, a rede neural pode ser usada para modelar o parâmetro de saída para condições desconhecidas de entrada (Ghatak, 2018). Na Figura 5 a seguir é mostrada uma arquitetura de um modelo MLP com algoritmo *backpropagation* onde x_n é o valor de entrada, y_m e z_{out} são os valores de saída das camadas oculta e de saída, respectivamente, w_{ji} e w_j são os pesos sinápticos, n e m são o número de neurônios nas camadas de entrada e oculta, respectivamente, e f é uma função de ativação.

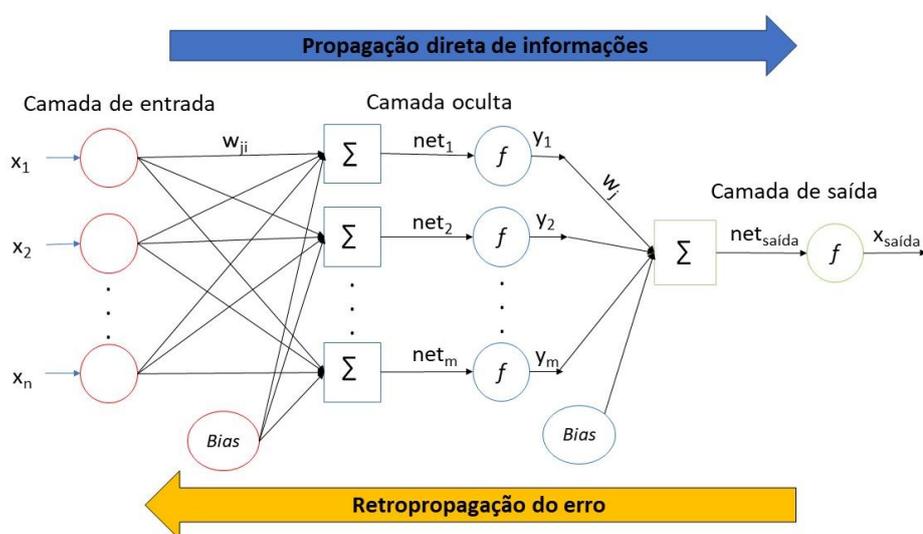


Figura 5 - Exemplo de funcionamento do algoritmo Backpropagation. Fonte: Elaboração própria.

Adaptado de Lee et al. (2020).

O treinamento da RNA visa encontrar uma função ótima que mapeie as entradas nas saídas desejadas. Uma rede bem treinada implica que ela aprendeu o modelo subjacente dos dados e pode generalizar bem para novos casos (Túpac, 2005). Para isso, os dados são comumente divididos em três conjuntos: treino, validação e teste. O conjunto de treino contém todos os dados representativos do domínio do problema e é usado na fase de treinamento para ajustar os pesos da rede. O conjunto de validação serve como um critério de parada potencial, avaliando ou confirmando a acurácia da melhor rede selecionada antes de implantá-la no sistema neural e/ou entregá-la ao usuário final. O conjunto de teste serve para verificar o desempenho da rede em relação a dados não vistos. (Basheer, 2000).

Um banco de dados é um conjunto de dados organizados em tabelas, registros e campos, que podem ser usados para desenvolver modelos de RNAs. Um banco de dados pode ser obtido de diversas fontes, como arquivos, bancos de dados relacionais, APIs, *web scraping*, sensores, etc. Um banco de dados pode conter dados estruturados, não estruturados ou semi-estruturados, dependendo do tipo e da forma dos dados.

As variáveis são as colunas do banco de dados que representam os atributos ou características dos dados. As variáveis podem ser divididas em: independentes (ou entradas) e dependentes (ou saídas). As variáveis independentes são os dados que alimentam a RNA, enquanto as variáveis dependentes são os dados que a RNA tenta prever ou classificar. As variáveis podem ser contínuas, binárias ou categóricas, dependendo do tipo e da escala dos dados (Microsoft, 2023).

O tamanho do banco de dados é o número de registros ou linhas do banco de dados que representam as amostras ou instâncias dos dados. O tamanho do banco de dados pode afetar o desempenho e a estabilidade da RNA, pois determina a quantidade e a qualidade dos dados disponíveis para o treinamento e o teste da RNA. Em geral, quanto maior o tamanho do banco de dados, melhor será a capacidade da RNA de aprender o padrão subjacente dos dados e generalizar para novos casos. No entanto, um tamanho muito grande do banco de dados também pode trazer alguns desafios, como maior tempo e custo computacional, maior risco de sobreajuste ou subajuste, maior necessidade de pré-processamento e validação dos dados, etc. (Túpac et al.,2005).

Os *outliers* são valores anormais ou extremos que se desviam significativamente da distribuição dos demais valores do banco de dados. Outliers podem ser causados por erros na coleta ou na medição dos dados, por variações naturais ou aleatórias dos dados, ou por eventos incomuns ou raros que afetam os dados. Outliers podem ter um impacto negativo no treinamento e no desempenho da RNA, pois podem distorcer a média, o desvio padrão, a correlação e outras medidas estatísticas dos dados, bem como introduzir ruído ou viés na RNA. Por isso, é importante identificar e tratar os outliers antes ou durante o treinamento da RNA. Algumas técnicas possíveis para lidar com outliers são (Ketha, 2019):

- Remover os outliers do banco de dados, se eles forem considerados erros ou anomalias.

- Substituir os outliers por valores médios, medianos ou modais do banco de dados, se eles forem considerados insignificantes ou irrelevantes.
- Transformar os outliers usando funções logarítmicas, raiz quadrada ou outras funções não lineares que reduzam a escala dos valores extremos.
- Usar métodos robustos que sejam menos sensíveis aos outliers, como redes neurais artificiais com funções de ativação não lineares, funções de perda robustas (por exemplo, erro absoluto médio) ou métodos estatísticos robustos (por exemplo, mediana ou intervalo interquartil).

2.5.3 Métricas de avaliação de desempenho

As métricas de avaliação de desempenho são fundamentais para a validação de modelos de previsão baseados em redes neurais artificiais (RNAs) aplicados em diferentes setores, incluindo o setor de biogás. O uso dessas métricas permite verificar a precisão das previsões realizadas pelos modelos, além de auxiliar na identificação de possíveis problemas no processo de treinamento das redes (Goodfellow, 2016). Uma das métricas amplamente utilizadas é a soma do erro quadrático (SSE), que é a soma dos quadrados das diferenças entre os valores observados e os valores previstos pela RNA. A equação da SSE pode ser escrita como na eq. (5):

$$SSE = \sum_{i=1}^n (Y - \hat{Y})^2 \quad (5)$$

Onde Y é o valor observado, \hat{Y} é o valor previsto pela RNA. Quanto mais próximo de 0 for o valor da SSE, melhor será o desempenho da RNA, pois isso indica que a rede está se aproximando mais dos valores observados.

Outra métrica muito utilizada é o coeficiente de determinação (R^2), que indica a proporção da variância total da variável de saída que é explicada pela RNA. O valor do R^2 varia entre 0 e 1, sendo que um valor próximo de 1 indica que a RNA é capaz de explicar a variabilidade da variável de saída. A equação do R^2 pode ser escrita como na eq. (6):

$$R^2 = 1 - \left(\frac{SSE}{SSR} \right) \quad (6)$$

Onde SSE (*Sum squared error*) é a soma do erro quadrático e SSR é a soma dos quadrados das diferenças entre os valores observados e a média dos valores observados.

No contexto do biogás, essas métricas podem ser aplicadas em modelos de previsão da produção de biogás a partir de diferentes substratos. Por exemplo, em um estudo realizado por Wang et al. (2021), foi utilizada uma RNA com a arquitetura MLP (*Multi-layer Perceptron*) para prever a produção de biogás a partir de diferentes misturas de resíduos agrícolas. A métrica SSE foi utilizada para avaliar o desempenho do modelo, e o valor obtido indicou que o modelo apresentou uma boa precisão na previsão da produção de biogás. Além disso, o valor do coeficiente de determinação (R^2) também foi reportado no estudo, mostrando que a RNA explicou a variabilidade da produção de biogás com precisão. Dessa forma, o uso de métricas de avaliação de desempenho é essencial para a validação e otimização de modelos de previsão baseados em RNAs, permitindo a sua aplicação em diferentes áreas.

2.6 Uso de redes neurais artificiais na temática do uso do biogás

As redes neurais artificiais têm sido bastante empregadas para a previsão de modelos. A RNA é uma poderosa ferramenta para modelagem de processos não lineares complexos e com múltiplas variáveis (Waewsak, 2010). O uso de ferramentas mais inteligentes para auxiliar a tomada de decisão em processos bioquímicos, por exemplo processos de digestão anaeróbia, estão cada vez mais sendo utilizadas. Nos últimos anos, a RNA tem sido aplicada com sucesso em bioprocessos não lineares multivariados como uma ferramenta útil para construir modelos e prever resultados com precisão para novos conjuntos de dados de entrada (Palaniswamy et al., 2016).

A revisão bibliográfica revela uma variedade de trabalhos que empregam modelos de RNA com distintas topologias para prever e otimizar biogás, conforme ilustra a Tabela 5. Na Tabela, para cada trabalho, são indicados os autores, a ferramenta utilizada, os parâmetros e variáveis do processo, a saída desejada e o número de casos, além de informações sobre a RNA como o algoritmo de treinamento, as funções de transferência/ativação, os números de entradas, o

número de camadas ocultas, o número de neurônios na camada oculta, o número de saídas, o R de treinamento, o R de teste, o R de validação e o R total. Como pode ser visto na tabela, dentre os exemplos de trabalhos citados, predomina o algoritmo de treinamento trainlm e as funções de ativação utilizadas são tansig, logsig e purelin. Contudo, são escassos os estudos que propõem modelos de RNA baseados em dados experimentais e da literatura como modelo geral, como é feito no presente trabalho.

Tabela 6 - Revisão da literatura de artigos que tratam de RNA aplicada à produção de biogás e biometano (continua)

Autor	Ferramenta utilizada	Parâmetros e variáveis	Saídas	Número de casos	Algoritmo de treinamento	Função de transferência	Nº de entradas	Nº de camadas ocultas	Nº de neurônios na camada oculta	Nº de saídas	R treinamento	R teste	R Validação	R Total
Qdais et al. (2009)	Rede Neural Artificial e Algoritmo Genético	Sólidos torais, sólidos voláteis, pH, temperatura e volume de biogás produzido.	Fração de metano (%CH ₄)	177	Trainlm	Tansig-Purilin	4	2	25	1	-	-	-	0,87
Gueguim et al. (2012)	RNA e Algoritmo Genético	Diferentes combinações em proporções percentuais diferentes de cada uma das matérias orgânicas testadas e o volume de biogás	Produção de biogás	125	Trainlm	Tansig-Purilin	5	1	2	1	-	-	-	0,9
Mahanty et al. (2013)	RNA e regressão	Importância relativa de cada lodo	Rendimento específico de metano	36	Trainlm	Tansig-Purilin	5	1	7	1	1	1	-	-
Jaroenpoj et al. (2014)	RNA e Algoritmo Genético	Importância relativa de cada substrato, pH, COD, VFA, VS	Fração metano, volume de biogás, nível de VS	130	Trainlm	Tansig-Purilin	5	1	8	3	0,99	0,98	-	0,99
Behera et al. (2015)	RNA e Algoritmo	Importância relativa de cada substrato, volume de biogás produzido	Fração de metano (%CH ₄)	131	Elm	Logsig-Logsig	2	1	15	1	-	0,88	-	-
Palaniswamy et al. (2016)	Rede neural artificial	Diferentes combinações de resto de comida, temperatura, pH e tempo de retenção hidráulica (TRH)	Produção de biogás	546	Trainlm	Tansig-Purilin	4	1	4 a 25	1	-	1	-	1

Tabela 5 – Revisão da literatura de artigos que tratam de RNA aplicada à produção de biogás e biometano (continua)

Autor	Ferramenta utilizada	Parâmetros e variáveis	Saídas	Número de casos	Algoritmo de treinamento	Função de transferência	Nº de entradas	Nº de camadas ocultas	Nº de neurônios na camada oculta	Nº de saídas	R treinamento	R teste	R Validação	R Total
Ghatak et al. (2018)	Rede neural artificial	Composição do substrato, temperatura e tempo de retenção hidráulica	Produção de biogás	4386	Trainlm	Logsig-Purilin	3	1	15	1	1	1	-	-
Beltramo et al. (2019)	RNA e Algoritmo Genético	Importância relativa de variáveis do processo	Vazão de biogás	100	Trainlm	Logsig-Purilin	11	1	10	1	0,96	0,96	0,95	-
Rego et al. (2019)	Rede neural artificial	Tipo de reator, temperatura do reator, pH, alcalinidade	Volume de biogás produzido	84	Trainlm	Logsig-Tansig-Purilin	4	2	10:10	1	-	-	-	0,88
Gonçalves Neto et al. (2020)	Rede neural artificial	Tipo de biomassa, tipo de reator/alimentação, VS, pH, OLR, TRH, temperatura e volume de reator	Produção de biogás	205	Trainlm – Trainbr – Traingdx	Tansig – logsig – Purilin	8	1	5 a 8	1	0,85	0,85	0,62	0,87
Mougari et al. (2021)	Rede neural artificial, algoritmo genético.	VS/TS, teor de carbono (C), relação carbono-nitrogênio (C/N) e tempo de digestão (DT)	Produção de biogás	2460	Trainlm	Logsig – Tansig	4	2	15 - 28	1	1	1	1	1
Chen et al. (2022)	Rede neural artificial	Qin, OLR, CODin, BODin, TSSin e pH	Remoção de COD, %CH ₄ no biogás e produção de metano	96	Trainlm	Tansig – Logsig	6	1	12	3	-	-	-	0,99

Fonte: Elaboração própria.

Gueguim (2012) e Qdais (2009) obtiveram resultados satisfatórios utilizando topologias semelhantes, alterando basicamente a quantidade de neurônios de cada camada ou o número de camadas da rede. Da mesma forma, Palaniswamy et al. (2016) conseguiram resultados convincentes configurando sua RNA como *multilayer perceptron* (MLP), com três camadas e com algoritmo de aprendizagem do tipo *Levenberg-Marquardt*. A RNA proposta foi capaz de prever o volume de gás produzido pelas amostras usadas no processo de validação, com uma média de 0,036% de erro de saída médio quadrático. Ainda foi possível verificar que uma faixa de pH entre 6 e 8 trazia melhores resultados de performance e também que temperaturas mais elevadas, dentro de uma faixa de 35 °C a 45 °C, melhoravam a produção de diferentes reatores. Para chegar ao modelo de RNA ideal, os autores realizaram o processo de treinamento por várias vezes, mas com diferentes configurações de camadas ocultas e quantidade de neurônios nestas camadas, podendo selecionar a combinação que melhor precisão tinha nas respostas de sua saída. A configuração que apresentou as melhores respostas ao final foi uma RNA do tipo *multilayer perceptron* de 3 camadas com configuração 4-17-1, correspondente as entradas, neurônios da camada oculta e saída, respectivamente, com algoritmo de aprendizagem *resilient backpropagation*.

Mougari et al. (2021) teve como objetivo desenvolver modelos de redes neurais artificiais (RNAs) globais bem estabelecidos, capazes de prever rendimento cumulativo de biogás (RCB) e rendimento cumulativo de metano (RCM) da digestão anaeróbica de diferentes substratos orgânicos sob condições operacionais ótimas. O efeito da razão sólido volátil para sólido total (VS/TS), teor de carbono (C), relação carbono-nitrogênio (C/N) e tempo de digestão (HRT) foram considerados. Foi aplicado algoritmo genético (AG) para otimizar a arquitetura da RNA. A arquitetura proposta foi uma configuração 4-2-1, indicando o número de entradas, número de camadas intermediárias e número de saídas da rede, respectivamente. Os resultados indicaram que a abordagem teve um bom desempenho na previsão de RCB e RCM e mostrou uma concordância decente com os dados experimentais, onde os valores de RMSE e R^2 para AG-RNA foram 0,0045, 0,0047 e 0,9996, 0,9998 para produção RCB e RCM, respectivamente.

Um modelo de RNA foi desenvolvida por Chen et al (2022) para prever os parâmetros de efluentes da digestão anaeróbica em um biorreator anaeróbio-

aeróbico integrado (IAAB) pré-comercializado. Um total de 6 parâmetros de entrada foram usados para prever a saída de 3 variáveis. O valor R obtido para remoção de COD (%), pureza de metano (%) e rendimento de metano foram 0,998, 0,983 e 0,990, respectivamente. A arquitetura de RNA usada para reação anaeróbica foi 6-12-3, indicando 6 entradas, 1 camada intermediária com 12 neurônios na camada oculta e 3 saídas. Foram testados 12 algoritmos de treinamento e o que apresentou melhor desempenho foi o algoritmo de treinamento Retropropagação de regularização bayesiana e a função de ativação tansig. Os modelos de RNAs desenvolvidas previram com sucesso as propriedades do efluente para a planta IAAB com precisão e erros mínimos.

Em Ghatak et al. (2018), o artigo demonstra a modelagem e otimização por RNA para previsão da produção de biogás utilizando esterco bovino como co-substrato, separadamente com pó de bambu, bagaço de cana-de-açúcar e serragem tanto para condições mesofílicas quanto para condições termofílicas. O estudo analisou o efeito de parâmetros de produção de biogás, como composição, temperatura e tempo de retenção hidráulica e obteve resultados com precisão de 99,7%.

Em Rego et al. (2019), o trabalho apresenta o desenvolvimento de Redes Neurais Artificiais (RNA) e Adaptive Neuro Fuzzy Inference System (ANFIS) para prever o volume de biogás. As variáveis do processo foram temperatura (°C), pH, relação FOS/TAC e tipo de biodigestores. Foi construído um banco de dados com as informações dos experimentos, dividindo-os em grupos de treinamento (67%) e teste (33%). O desempenho dos modelos neurais foi avaliado pelo coeficiente de determinação (R^2) e erro índice (RMSE). O modelo desenvolvido a partir da modelagem RNA e ANFIS foi satisfatório, atingindo um valor de R^2 de 0,881.

Em Gonçalves Neto et al. (2020), o estudo desenvolve um modelo matemático através de redes neurais artificiais (RNA), a partir dos 8 parâmetros de interesse listados na tabela como parâmetros de entrada e analisa a produção acumulada de biogás como parâmetro de saída. O modelo encontrado apresentou valores aceitáveis de coeficiente de determinação (valores entre 0,6167 e 0,9929). No estudo, foi descoberto que a produção de biogás para as três biomassas de interesse foi maior quando as temperaturas estão dentro da faixa de condições termofílicas, com um máximo local para condições mesofílicas. Biodigestão

otimizada de CD ou FW permite maior teor de VS (cerca de 15-20%) do que para FVW (inferior a 10%). Observou-se também que a biodigestão do FVW levou a maior produção cumulativa de biogás (cerca de duas vezes o valor encontrado para FW e CD).

3 Metodologia

Um diagrama contendo as etapas a serem seguidas e implementadas para o desenvolvimento de uma RNA com capacidade para previsão de produção acumulada de biogás (PAB), produção acumulada de metano (PAM) e também da porcentagem de metano ($\%CH_4$) presente nas amostras, bem como a avaliação de melhorias visando um melhor desempenho, é apresentado na Figura 6.

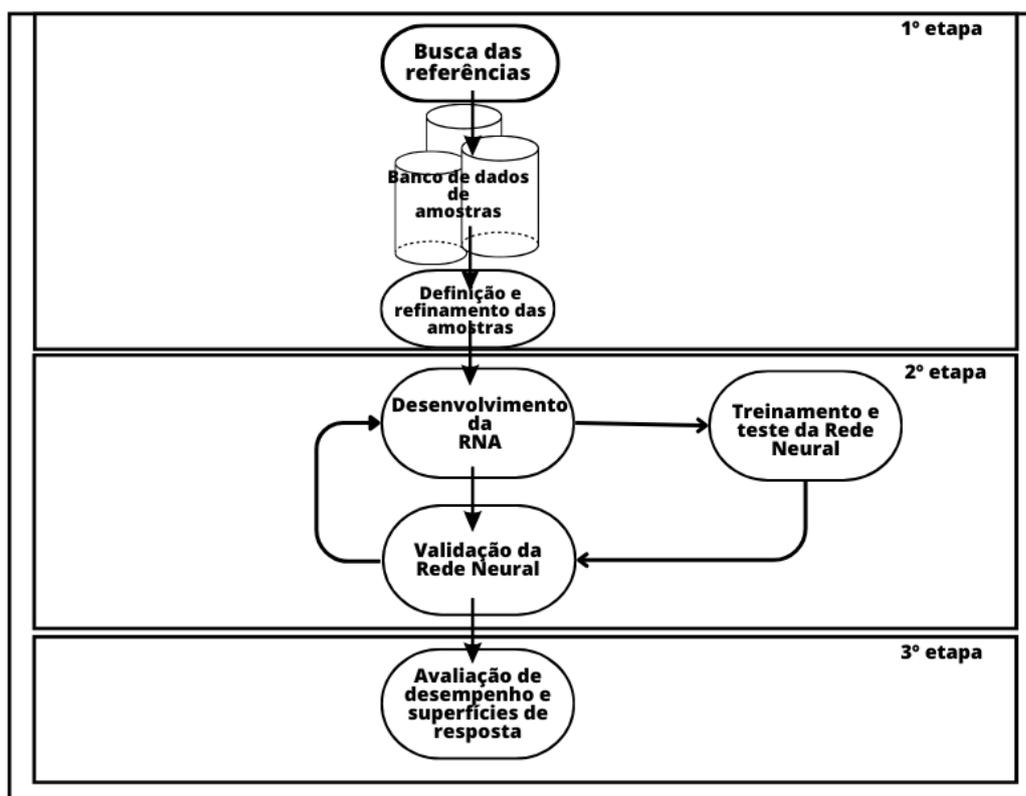


Figura 6 - Diagrama com as etapas da metodologia. Fonte: Elaborada pelo autor

O diagrama apresentado na Figura 11 representa o fluxo de trabalho utilizado para a realização deste estudo. Primeiramente, um banco de dados foi compilado com informações relacionadas aos parâmetros pré-definidos para a produção acumulada de biogás, acumulada de metano e porcentagem de CH_4 . Esse banco de dados foi então utilizado para a seleção de amostras que seriam utilizadas no processo de treinamento e teste das redes neurais artificiais (RNAs). Após a seleção das amostras, o processo de treinamento e teste foi realizado em três diferentes topologias de RNA para cada um dos bancos de dados de interesse.

Posteriormente, foi realizada uma etapa de validação, onde os resultados obtidos foram comparados com os dados do próprio banco de dados que não foram utilizados nos processos de treinamento e teste. Esse processo de validação tem como objetivo verificar a eficácia e convergência da rede desenvolvida, ou seja, sua capacidade de generalização dos resultados quando submetida a entradas ainda não conhecidas.

A terceira etapa do fluxo de trabalho consistiu na avaliação dos resultados e desempenho das RNAs, bem como na geração e avaliação de diferentes superfícies de resposta combinando as variáveis de interesse. Para apresentar os resultados, foi seguida uma sequência iniciando pela limpeza e refinamento dos bancos de dados, seguida pelos resultados obtidos na etapa de treinamento, teste e validação das três diferentes topologias implementadas e posteriormente apresentando os resultados obtidos nas etapas de confecção das superfícies de resposta. Esses resultados são importantes para compreender melhor o processo de produção de biogás e metano, bem como para apontar possíveis melhorias no processo e sugestões para futuras pesquisas.

3.1 Definição da base de dados com todas as biomassas e refinamento das amostras

Para a confecção do banco de dados completo, foi necessário um processo rigoroso de seleção e análise de artigos, bem como uma revisão cuidadosa dos dados coletados para garantir a precisão e confiabilidade das informações utilizadas no estudo. Além disso, foi necessário realizar a padronização dos dados e a transformação das unidades de medida em uma escala comum, a fim de possibilitar a utilização desses dados em uma análise integrada. Todos esses esforços foram necessários para garantir a validade e robustez do estudo realizado. Foi realizado um trabalho exaustivo de pesquisa bibliográfica para a obtenção de artigos científicos que contivessem dados suficientes para o preenchimento dos parâmetros pré-definidos. Para o presente trabalho, foram analisados 47 trabalhos científicos publicados entre 2003 e 2020 ao redor do mundo, incluindo Garcia-Peña et al. (2011), Yong et al. (2015), Chen et al. (2014), Wang et al. (2014), Sheng et al. (2013), Kawai et al. (2014), Deepanraj (2014), Zhang et al. (2013), Tampio et al.

(2014), Li and Jin (2015), Kim et al. (2006), Masebinu et al. (2018), Nagao et al. (2012), Ganesh et al. (2014), Shen et al. (2013), Ventura et al. (2014), Wu et al. (2016), Banks et al. (2010), Wu et al. (2015), Li and Jin (2018), Xiao (2018), Micolucci et al. (2018), Gottardo (2018), Xiao and Qin (2018), De Gioannis (2017), Bouallagui (2004), Grimberg (2015), Hobbs (2015), Jabeen (2015), Kuczman (2018), Paudel (2017), Chu (2008), Ventura (2014), Bouallagui (2003), Browne and Murphy (2013), El-Mashad (2010), Wang and Zhao (2009), Ariunbaatar (2015), Koyabashi (2012), Lee (2010), Micolucci (2014), Vats et al. (2019), Zhang et al. (2019), Shamurad (2020), e Rico (2020).

Considerando a complexidade da digestão anaeróbia, que envolve diversos parâmetros experimentais e fatores biológico-físico-químicos, o primeiro passo para o cumprimento da proposta do trabalho consistiu em definir quais propriedades seriam relevantes para a coleta de dados. O critério estabelecido foi a importância de cada parâmetro para o processo de biodigestão, como foi discutido na revisão bibliográfica, e a disponibilidade de dados nos artigos analisados. Essas propriedades relevantes estão listadas na Tabela 6. Alguns artigos disponibilizavam dados numéricos em gráficos, que foram lidos e extraídos com o auxílio da ferramenta Pega Ponto versão 1.0.

Tabela 7 - Propriedades selecionadas no levantamento de dados. Fonte: Elaboração própria

Referência
País
Tipo de substrato
Configuração da AD (Regime de alimentação e número de estágios)
TS (sólidos totais, %)
VS (sólidos voláteis, %)
VS/TS
OLR (g VS/l.d)
HRT (d)
T (°C)
Volume do biodigestor (m ³)
Produção de biogás (l/g.VS)
Produção de metano (l/g.VS)

A seleção dos 47 trabalhos científicos mencionados acima foi precedida por uma revisão bibliográfica mais abrangente, na qual algumas referências foram

excluídas para o presente trabalho, em virtude de não contemplarem muitas das propriedades indicadas como relevantes para a metodologia, ou de não fornecerem informações suficientes sobre algumas dessas propriedades. Após a seleção, alguns artigos permitiram a análise de mais de um cenário, pois em determinados casos, a codigestão de substratos era efetuada em diferentes níveis de concentração, e em outros, alterava-se a taxa orgânica para um mesmo tipo de resíduo. Convém mencionar que, para possibilitar a comparação integral de todos os casos em relação a todas as suas propriedades, os casos de modo de alimentação batelada receberam uma carga orgânica (OLR) artificial. Esta foi inferida a partir do tempo de retenção (HRT) do processo. Já em algumas situações em que ocorria a codigestão de distintas matérias orgânicas, foi necessário realizar o cálculo para se obter o teor de TS e VS do substrato a ser introduzido no sistema, conforme a proporção indicada pelo próprio artigo. Ademais desses cálculos iniciais, também foi necessário realizar frequentemente o trabalho de conversão de unidades.

Assim, foi elaborado um banco de dados consistente com parâmetros de interesse para o trabalho, que são: tipos de substratos, tipo de biodigestor, TS, VS, TS/VS, pH, OLR, HRT, temperatura e volume do biodigestor, produção de biogás, produção de metano e concentração de CH₄. O tipo de substrato é descrito como uma variável discreta de modo a distinguir FW (definida como biomassa processada de cozinhas domésticas e comerciais), FVW (definida como biomassa não processada das fontes citadas anteriormente, do mercado e distribuidores de frutas e vegetais) e CD (apenas entre FW e FVW). Outra variável discreta é definida para representar simultaneamente o tipo de reator/alimentação (B, C, SC) e o número de estágio do digestor (um, dois ou três estágios). Depois, foram definidas as 6 variáveis contínuas de interesse relacionadas a biomassa e ao processo de digestão anaeróbica, são elas: VS (em %), pH no biodigestor, OLR (em g.VS/ (l.d)), HRT (em dias, calculado para toda a biodigestão), temperatura (em °C) e volume do reator (em m³). A seguir, apresenta-se a Tabela 7 para correlacionar as variáveis discretas.

Tabela 8 - Tabela com as variáveis discretas. Fonte: Elaboração própria

Variáveis discretas					
Substrat	Tipo	Valor	Biodigestor	Tipo	Valor
		FVW		-1	
	FW	0		Batelada 2 estágios	-2
	CD	1		Semi contínuo 1 estágios	-1
				Semi contínuo 2 estágios	0
				Semi contínuo 3 estágios	1
				Contínuo 2 estágios	2
				Contínuo 1 estágio	3

Para cada artigo científico analisado, os parâmetros pré-definidos foram identificados e as informações correspondentes foram extraídas e organizadas de maneira uniforme. Essas informações foram então armazenadas em um banco de dados estruturado, no qual cada parâmetro foi associado a um valor numérico ou categórico, com base nas informações apresentadas nos artigos. Esse banco de dados foi utilizado para ajustar um modelo de rede neural artificial (RNA), o qual permitiu o treinamento, teste e validação da rede para previsão da produção acumulada de biogás, acumulada de metano e a concentração de CH₄ em biodigestores anaeróbicos. Para que a RNA pudesse acessar as informações armazenadas, os dados foram organizados como vetores e matrizes, de acordo com a especificação da rede utilizada.

Após a criação inicial do banco de dados, contendo 2098 amostras, os dados foram divididos em 3 diferentes conjuntos, cada um correspondendo a uma resposta específica a ser analisada: produção acumulada de biogás, produção acumulada de metano e concentração de CH₄. O tratamento dos dados consistiu em eliminar os valores discrepantes, normalizar os dados e verificar a distribuição dos mesmos. A análise estatística dos dados foi realizada por meio de histogramas gerados no Excel para cada parâmetro de interesse, que mostram a frequência relativa de cada valor ou intervalo de valores. Os histogramas permitem visualizar a forma, o centro e a dispersão dos dados, bem como identificar possíveis tendências ou padrões. Cada um dos bancos de dados (biogás, metano e %CH₄) foi dividido em três grupos, sendo 60% dos dados utilizados para o treinamento e desenvolvimento da rede, enquanto os outros 40% foram utilizados nos testes funcionais das redes e na validação. Dentro desses 40%, foram separados 20% para teste e 20% para

validação. Essa separação aleatória foi realizada utilizando a função *dividerand* do software Matlab. O objetivo dessa divisão foi evitar o *overfitting* da rede neural, ou seja, garantir que a rede fosse capaz de generalizar o aprendizado para novos dados, além de avaliar a capacidade preditiva da RNA em relação aos dados de teste e validação. Detalhes específicos sobre a formatação de cada grupo de dados serão apresentados ao longo dos subcapítulos seguintes, a fim de fornecer informações mais precisas sobre o processo de treinamento, teste e validação da RNA.

3.2 Definição do modelo aplicado a estimação e treinamento, validação e teste da RNA

Esse trabalho foi baseado no trabalho de Gonçalves Neto et al. (2021). Foram necessárias modificações de modo a adequar a topologia à disponibilidade de dados e características do problema proposto. Tendo como base as referências e a disponibilidade de dados e seus respectivos parâmetros, o projeto de implementação deste trabalho contemplou o desenvolvimento de um algoritmo baseado exclusivamente em modelos de RNA que tivessem capacidade de prever a produção acumulada de biogás, a produção acumulada de metano e a sua porcentagem em um sistema de biodigestão, utilizando como dados de entrada os dados levantados na etapa anterior. As RNAs, com sua arquitetura MLP, foram utilizadas para análise das faixas das variáveis e obtenção de superfícies de resposta, intuindo a região. A metodologia utilizada neste estudo consistiu na implementação de redes neurais artificiais (RNAs) compostas por três ou mais camadas interligadas de neurônios artificiais.

A determinação da quantidade de neurônios na camada oculta foi realizada com base em dois critérios: o critério de erro e um número constante de parâmetros efetivos. Para o presente trabalho, o número de neurônios na camada oculta variou de 8 a 11 para o treinamento. Esse intervalo foi definido depois de testes indicarem que mais do que 11 neurônios e menos que 8 na camada oculta não resultavam em melhora significativa do desempenho da rede, além de ter um custo maior de treinamento (tempo gasto) à medida que o número de neurônios na camada oculta aumenta. Além disso, foram utilizadas as funções de ativação tangente hiperbólica (*tansig*) e sigmóide logística (*logsig*) nas camadas ocultas, enquanto a camada de

saída utilizou uma função de transferência linear (purelin). Foram adotados três algoritmos de treinamento para as RNAs: traingdx, trainlm e trainbr. O algoritmo traingdx utiliza a abordagem gradiente descendente, que emprega o gradiente negativo para minimizar o erro. Já os algoritmos trainlm e trainbr utilizam o método Quasi-Newton para minimizar o erro, com a diferença de que o trainlm emprega diretamente o Método de Levenberg-Marquardt, enquanto o trainbr utiliza Regularização Bayesiana. A escolha desses algoritmos levou em consideração as suas características e capacidade de desempenho em relação à natureza do problema de modelagem em questão. O treinamento, teste e validação foram realizados no MATLAB 2019b usando o Deep Learning Toolbox para otimizar pesos e valores de bias.

Para determinar a qualidade do treinamento, o valor da soma do erro quadrado (SSE) é utilizado. A acurácia preditiva do modelo foi estimada pelo coeficiente de determinação (R^2). A Tabela 8 a seguir apresenta um resumo das condições para a definição do modelo da RNA.

Tabela 9 - Condições para definição do modelo da RNA. Fonte: Elaboração própria

Número de neurônios da camada de entrada	8
Número de neurônios da camada de saída	1
Número de neurônios da camada oculta	(8,9,10,11)
Funções de ativação da camada oculta	Tansig, logsig
Algoritmos de treinamento	Traingdx, trainbr, trainlm
Função da camada de saída	Purelin

Como critério de parada, o algoritmo utiliza o critério que primeiro for alcançado, podendo ser o número máximo de iterações necessárias para atingir a melhor performance ou o número de tentativas para se atingir a melhor performance. Uma vez atingido algum dos limites estabelecidos, a melhor performance é armazenada bem como os pesos sinápticos e demais parâmetros da rede para posterior utilização nos testes de funcionalidade, validação e avaliação de performance. Após a definição da melhor performance da rede neural, essa topologia é utilizada para a análise dos resultados e a obtenção de superfícies de

respostas de modo a identificar a influência de cada parâmetro de entrada nas respostas desejadas.

A melhor performance refere-se aos resultados mais precisos e eficientes obtidos pelas redes neurais artificiais utilizadas no estudo. Essa avaliação de desempenho é realizada para mostrar que as estimações e predições dos modelos tendem a se aproximar dos resultados reais coletados a partir da literatura, evidenciando a capacidade de estimação das RNAs. Os valores obtidos pelas saídas das redes são apresentados em forma de gráficos para melhor compreensão dos resultados.

3.3 Geração das superfícies de resposta

Para visualizar as relações entre as variáveis independentes (parâmetros operacionais) e a variável dependente (produção de biogás ou metano), foram geradas superfícies de resposta para cada um dos modelos obtidos. As superfícies de resposta são gráficos tridimensionais que mostram como as variáveis se influenciam mutuamente. Os resultados foram comparados com a literatura existente e as ordens de grandeza esperadas para cada variável, conduzindo-se uma análise minuciosa. Foram identificados padrões consistentes e discrepâncias notáveis, que foram comentados e explicados com base nos princípios físicos e químicos envolvidos no processo de biodigestão anaeróbia. Isso permitiu compreender melhor os processos de produção de biogás e metano, contextualizando-se as descobertas no âmbito da literatura científica e das expectativas teóricas.

As superfícies de resposta foram geradas utilizando o software Matlab. O código do programa desenvolvido carrega a rede neural com melhor desempenho definida nas etapas anteriores, juntamente com o banco de dados de treinamento. As superfícies de interesse foram definidas a fim de avaliar a influência de duas variáveis de entrada combinadas na saída de interesse. As variáveis que não eram de interesse para a plotagem das superfícies foram definidas com valores fixos, sendo que esses valores foram determinados com base na média dos valores para a classe de maior representação para cada variável, de acordo com os histogramas apresentados. A escolha de determinar os valores fixos com base na média dos

valores para a classe de maior representação das variáveis é justificada estatisticamente. Isso garante que os valores fixos estejam próximos do comportamento mais comum das variáveis, tornando as superfícies de resposta mais representativas da realidade. O domínio de cada superfície foi definido como o valor mínimo e máximo de cada variável de entrada.

Vale ressaltar que a utilização de RNAs e análise estatística para a geração de superfícies de resposta tem sido amplamente empregada em diversas áreas de pesquisa, incluindo a produção de biogás. Um exemplo é o estudo de Zhang et al. (2020), que utilizaram uma abordagem semelhante para avaliar a influência de diferentes variáveis de processo na produção de biogás a partir de resíduos alimentares. Outra referência relevante é o trabalho de Tchakouté et al. (2018), que utilizaram RNAs e análise estatística para modelar a produção de biogás a partir de substratos diferentes, com o objetivo de identificar as condições ótimas de operação para cada um dos substratos.

Em suma, a metodologia adotada para a geração das superfícies de resposta foi cuidadosamente planejada e baseada em técnicas estatísticas e de RNAs amplamente utilizadas em diversas áreas de pesquisa, incluindo a produção de biogás. As superfícies geradas permitiram a identificação de pontos ótimos de operação para cada um dos sistemas avaliados, o que pode contribuir significativamente para o desenvolvimento de processos mais eficientes e sustentáveis de produção de biogás.

3.4 Desenvolvimento e avaliação de modelos de RNA específicos por biomassa

A partir dos resultados das superfícies de resposta para os bancos de dados completos, foram identificadas algumas limitações e inconsistências, o que motivou a necessidade de uma análise mais aprofundada. Reduziu-se o banco de dados em subconjuntos específicos com base nas categorias de biomassa. Isso permitiu desenvolver modelos especializados que consideram as nuances individuais de cada tipo de biomassa. Realizou-se uma análise estatística comparativa entre esses modelos especializados e o modelo global. Isso incluiu a comparação das previsões, bem como a identificação das vantagens e desvantagens de utilizar conjuntos de

dados mais especializados. Essa abordagem refinada proporcionou uma visão mais precisa e detalhada das propriedades de produção de biogás e metano de diferentes tipos de biomassas. Da mesma forma, aplicou-se a geração de superfícies de resposta aos subconjuntos de biomassa, mantendo os valores fixos globais para que pudesse ser feita uma comparação entre os modelos. Realizou-se uma comparação metódica entre essas superfícies de resposta especializadas, a literatura relevante e as ordens de grandeza esperadas com os modelos obtidos do banco de dados global. Os comentários detalhados destacaram os padrões coerentes e incoerentes encontrados. Além disso, analisaram-se as implicações dessas descobertas, explorando como essas variações podem ser aproveitadas para otimizar o processo de produção de biogás e metano.

Essa abordagem sistemática e minuciosa, combinada com a especialização dos dados, permitiu uma compreensão aprofundada dos processos de produção de biogás e metano, fornecendo informações valiosas para a pesquisa futura e a aplicação prática no campo da energia renovável.

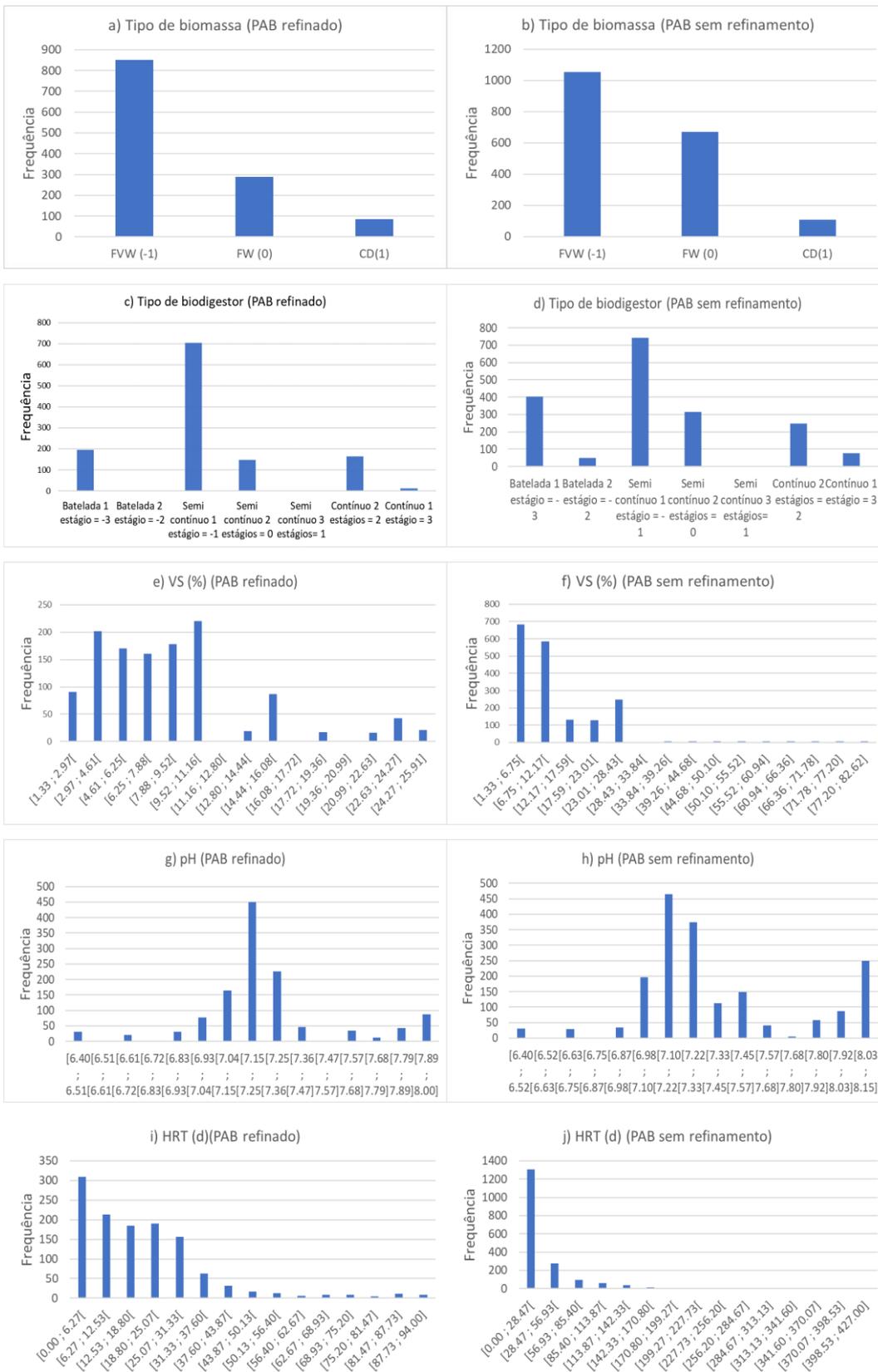
4 Resultados e Discussões

Nessa seção, são apresentados os resultados obtidos a partir da metodologia proposta e da implementação das redes neurais.

4.1 Consolidação dos bancos de dados completos

Como descrito na seção de metodologia, após o trabalho de coleta de dados da literatura, foi elaborado um banco de dados geral com 2098 amostras contendo valores das 8 variáveis de entrada pré definidas e as 3 variáveis de saída estudadas. Esse banco de dados foi posteriormente dividido em 3 diferentes bancos de dados de acordo com a resposta desejada. Primeiramente, foram excluídos desses bancos de dados de resposta linhas que continham dados incompletos, visto que no banco de dados geral algumas amostras não continham todos os valores de entrada que foram definidos para o presente trabalho. Com essa etapa inicial, cada banco de dados formado passou por uma etapa de análise estatística para localizar valores muito discrepantes e que estariam fora da faixa de interesse para a etapa posterior do estudo que é a construção das superfícies de resposta. Para isso, foram construídos histogramas para as análises visando analisar o espalhamento dos dados de cada variável e de cada resposta.

Para o banco de dados referente a produção acumulada de biogás (PAB), um conjunto de 1832 amostras foram obtidas ao final dessa etapa inicial (PAB sem refinamento). Os dados que foram considerados muito discrepantes após a análise dos histogramas foram cortados no banco de dados, formando um segundo grupo de dados (PAB refinado). Os histogramas para os dois bancos de dados PAB (sem refinamento e refinado) são mostrados na Figura 7 a seguir.



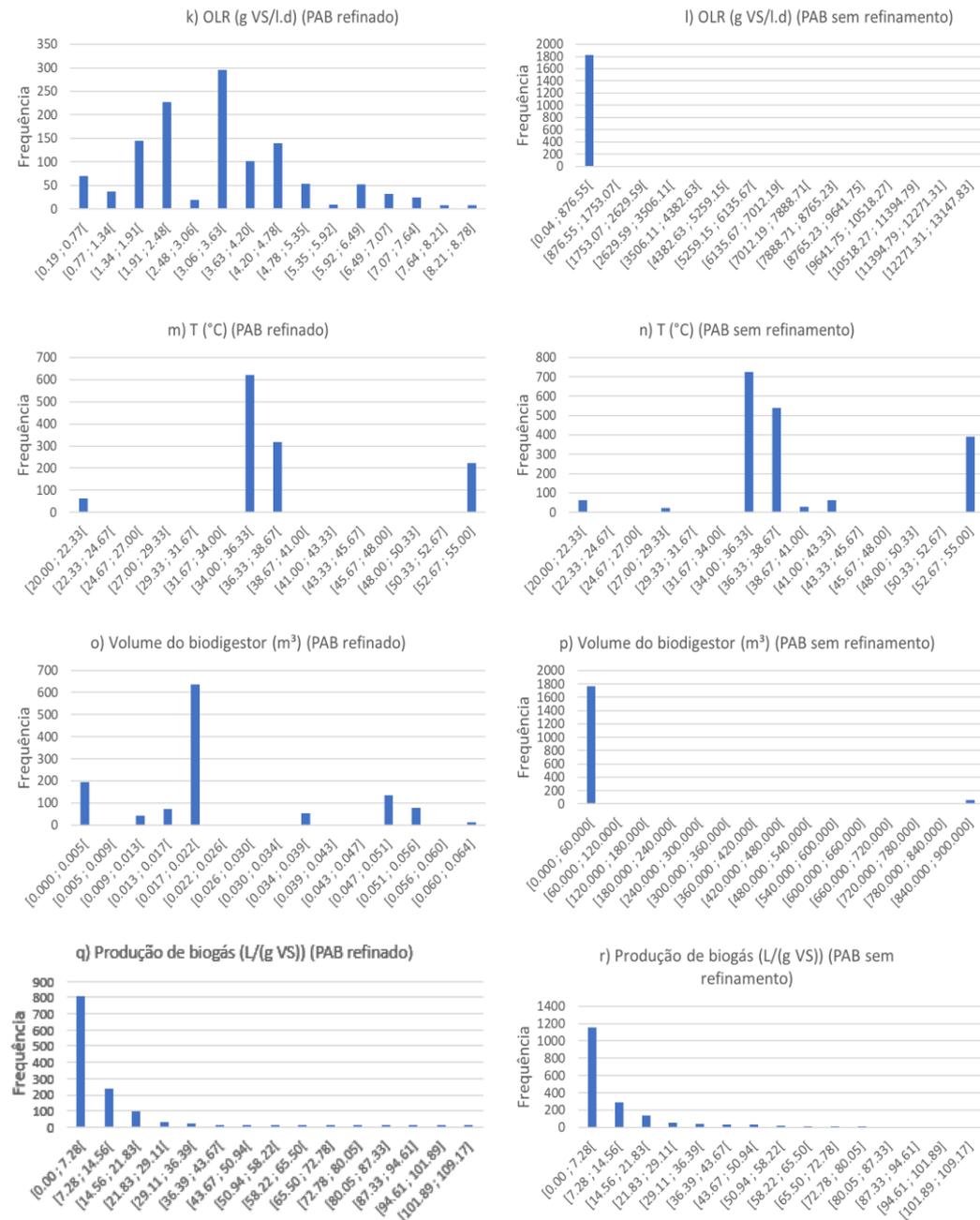


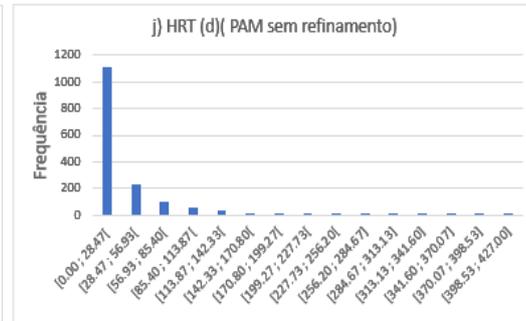
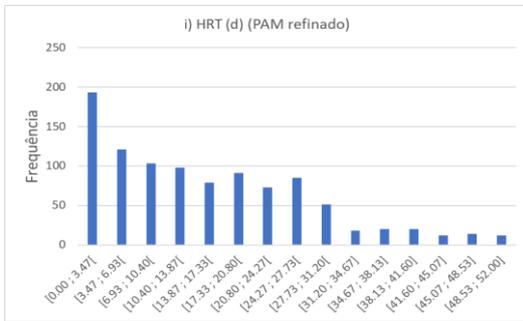
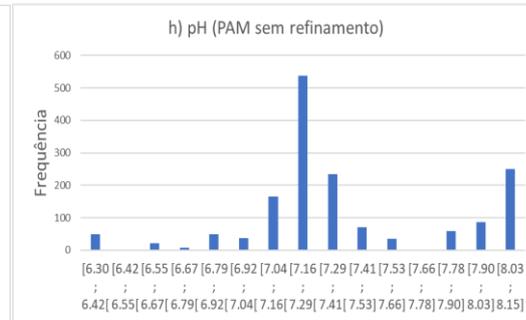
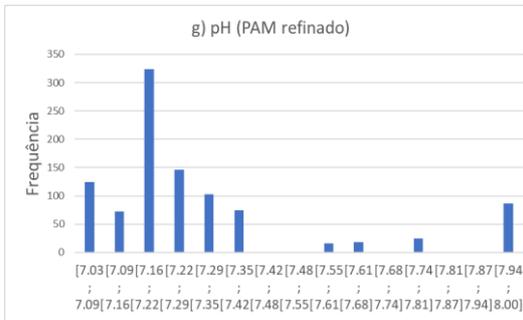
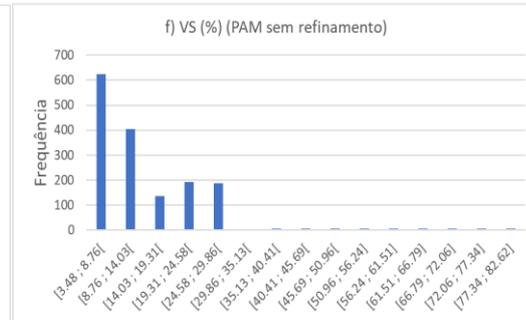
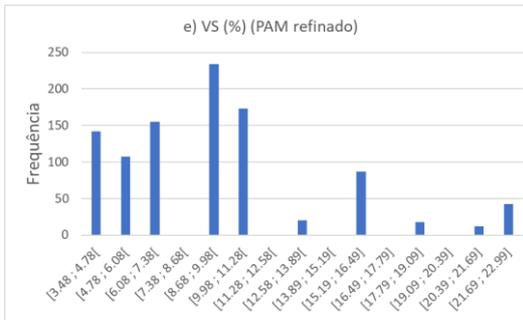
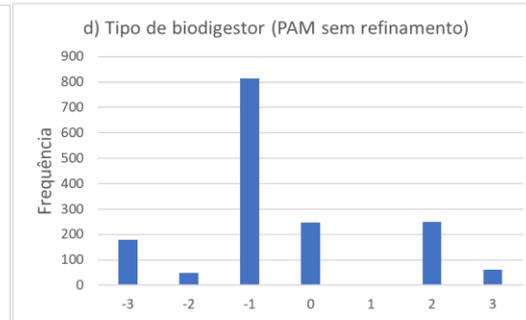
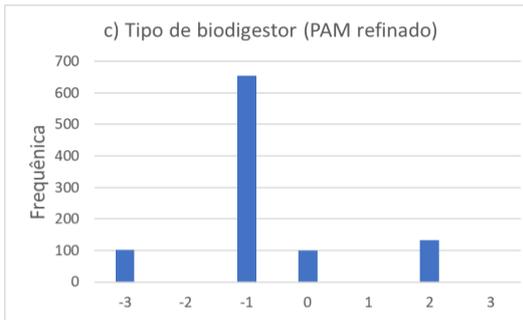
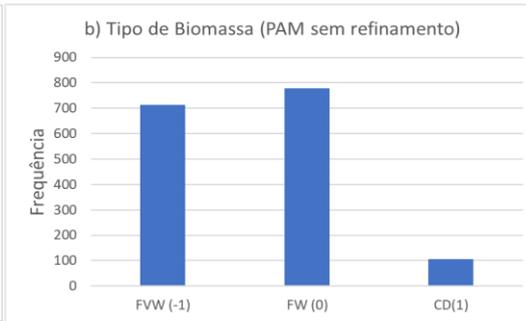
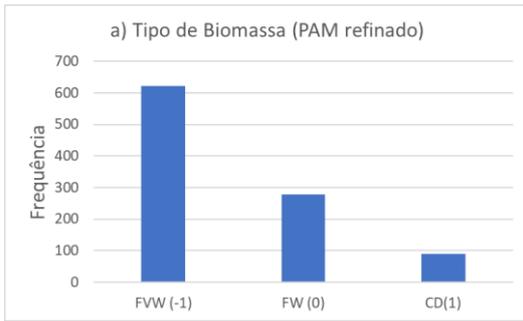
Figura 7 - Histogramas para os bancos de dados PAB: comparação das distribuições de “PAB sem refinamento” e “PAB refinado” para tipo de biomassa, tipo de reator, teor de sólidos voláteis (VS), pH, taxa de carga orgânica (OLR), tempo de retenção hidráulica (HRT), temperatura, volume do reator e produção acumulada de biogás

Os histogramas gerados a partir dos dados sem refinamento mostraram que existem regiões do banco de dados com pouca ocorrência de valores para algumas variáveis. Em particular, as regiões com tipo de reator semicontínuo de 3 estágios, valores de VS acima de 28,43%, HRT acima de 142,33 dias e volume de biodigestor

acima de 60 m^3 foram pouco populadas. Já para a variável OLR, os dados se concentraram principalmente na primeira classe do histograma. Para o histograma de produção de biogás, é possível verificar que existe uma proporção de cerca de dois terços na primeira classe e um terço nas demais. O refinamento contemplou a retirada de dados de OLR acima de $8,78 \text{ gVS/ld}$ (valores muito altos foram introduzidos no cálculo de um pseudo-OLR para reator tipo batelada), HRT acima de 94 dias, VS acima de 28,43% e volumes de reator acima de $0,1 \text{ m}^3$. Essa etapa levou ao banco PAB refinado com 1224 amostras, 66,81% do conjunto de dados do banco PAB sem refinamento.

Olhando para a distribuição do conjunto de dados menor, pode-se observar que a condição de codigestão está pouco representada, assim como tipos de reator batelada de 2 estágios, semi contínuo de 3 estágios e contínuo de 1 estágio. Com relação ao teor de VS, os dados coletados estão apresentando uma distribuição quase homogênea na faixa de 2,97% a 11,16%, enquanto o pH segue uma distribuição quase normal centrada na faixa de 7,15-7,25. Os valores de HRT concentram-se em uma duração de até 1 mês, conforme relatado na literatura. A faixa de OLR com maior população é $3,06\text{-}3,63 \text{ gVS/(l.d)}$, mas valores tão altos quanto $8,78 \text{ gVS/(l.d)}$ podem ser encontrados. A distribuição da temperatura é bimodal, com a maioria dos dados em torno de 35°C (temperatura mesofílica) e outros próximos de 55°C (temperatura termofílica). Reatores com volume de aproximadamente 20L são os mais representados, confirmando que os dados são coletados em escala laboratorial ou piloto. A PAB mostrou que o refinamento do conjunto de dados não afetou significativamente o histograma, pois os valores mais baixos ainda são os mais abundantes e as classes com valores mais altos ainda foram representadas.

O banco de dados referente à produção acumulada de metano (PAM) foi obtido após a etapa inicial, contendo 1598 amostras (PAM sem refinamento). Os dados que apresentaram valores muito discrepantes, conforme a análise dos histogramas, foram excluídos do banco de dados, originando um segundo conjunto de dados (PAM refinado). A Figura 8 mostra os histogramas para os dois conjuntos de dados PAM (sem refinamento e refinado).



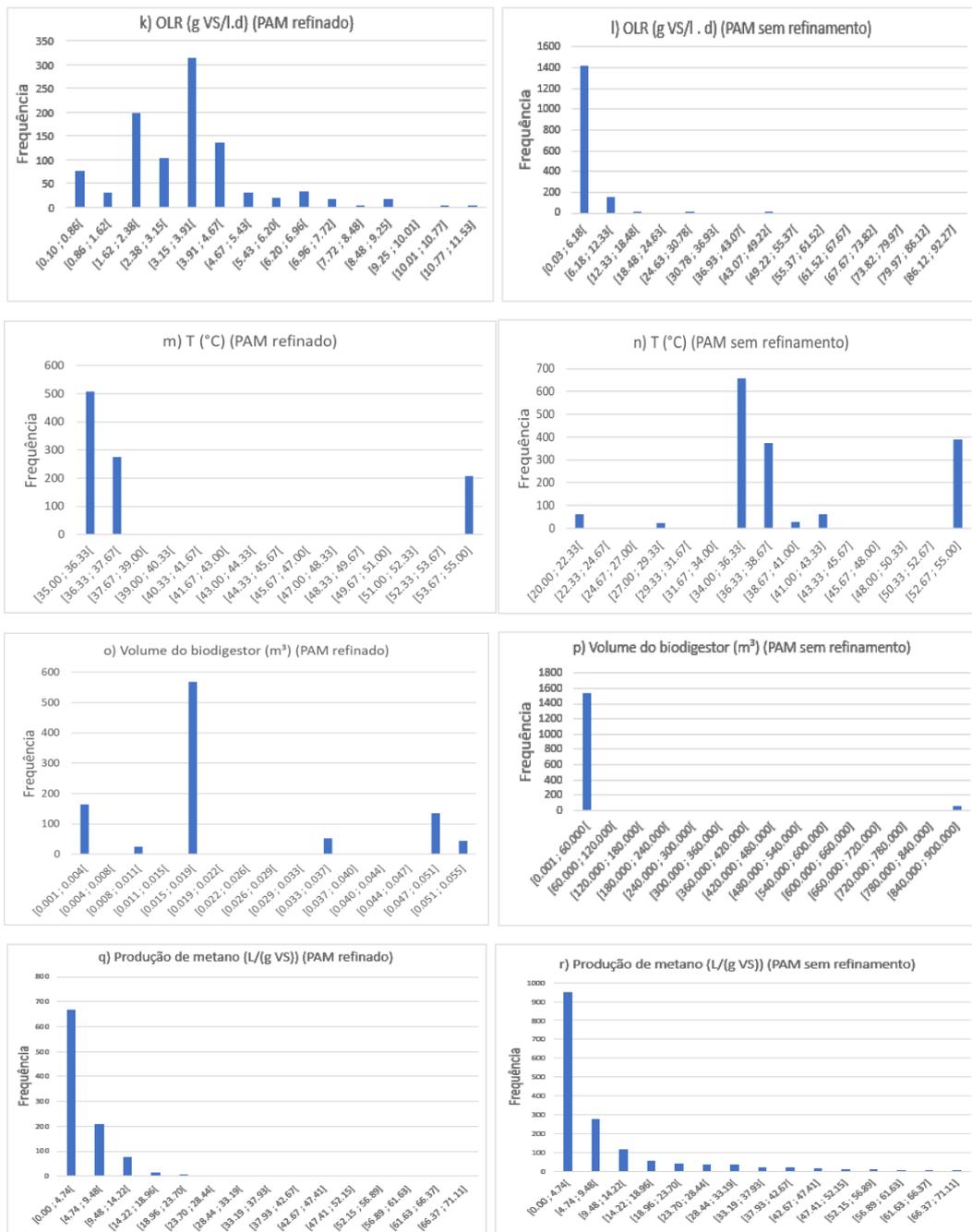


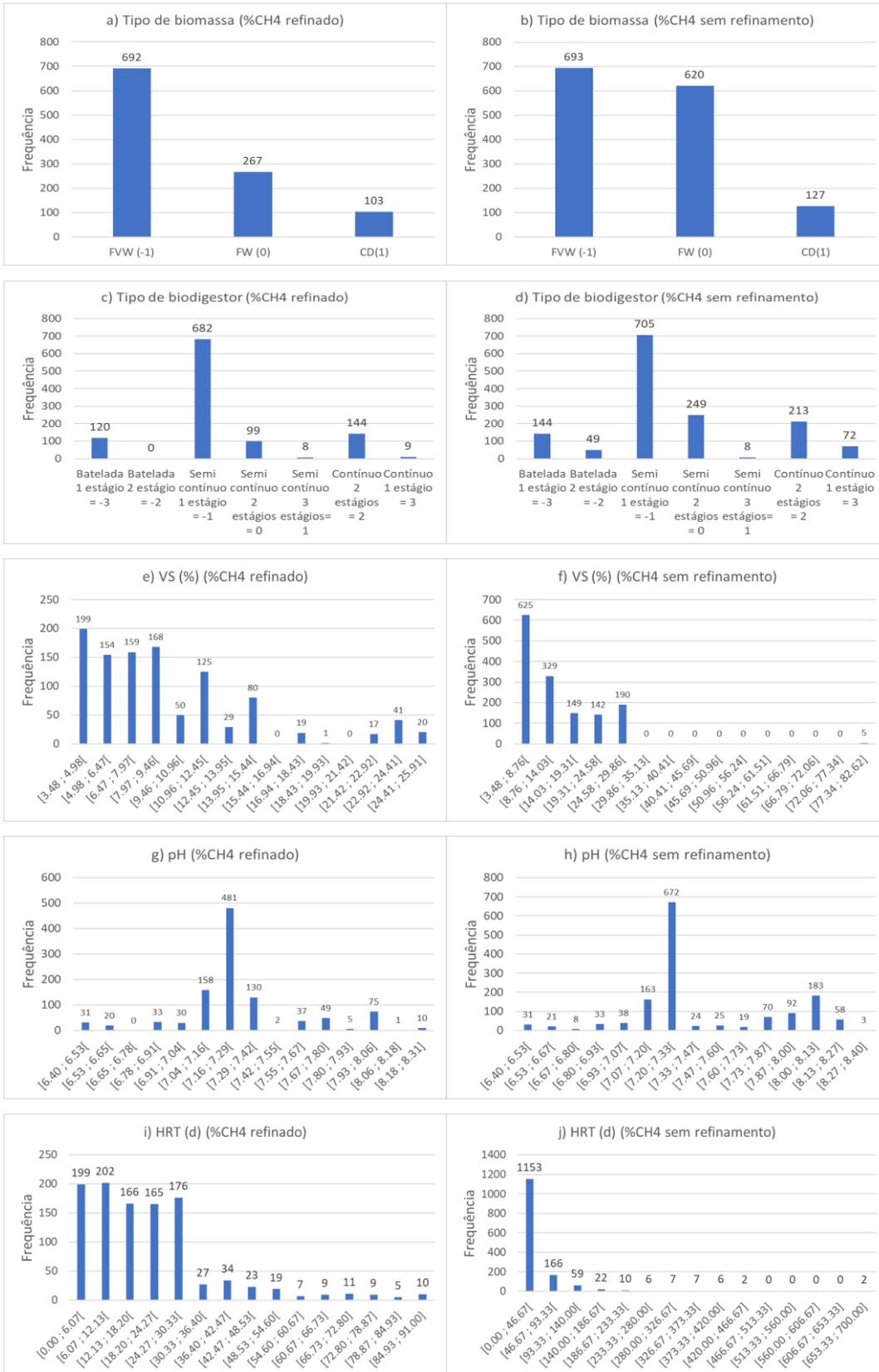
Figura 8 - Histogramas para os bancos de dados PAM: comparação das distribuições de “PAM sem refinamento” e “PAM refinado” para tipo de biomassa, tipo de reator, teor de sólidos voláteis (VS), pH, taxa de carga orgânica (OLR), tempo de retenção hidráulica (HRT), temperatura, volume do reator e produção acumulada de metano

A partir dos dados sem refinamento, os histogramas indicaram que algumas variáveis apresentaram regiões com baixa frequência de valores no banco de dados. Em especial, as regiões com reator semi contínuo de 3 estágios, VS superior a 29,86%, HRT superior a 113,87 dias e volume de biodigestor superior a 60 m³

foram pouco representativas. Quanto à variável OLR, os valores se concentraram principalmente na primeira classe do histograma. Para o histograma de produção de biogás, verificou-se que cerca de dois terços dos valores estavam na primeira classe e um terço nas demais. O refinamento consistiu na exclusão de dados de OLR superior a 12,36 gVS/l.d (valores muito elevados foram calculados para um pseudo-OLR para reator tipo batelada), HRT superior a 54,11 dias, VS superior a 29,86%, pH abaixo de 7,04 e volumes de reator superiores a 0,1 m³. Essa etapa resultou no banco PAM refinado com 990 amostras, 61,95% do conjunto de dados do banco PAM sem refinamento.

A distribuição do conjunto PAM refinado mostrou que a codigestão foi pouco representativa, assim como os reatores batelada de 2 estágios, semi contínuo de 3 estágios e contínuo de 1 estágio. Quanto ao teor de VS, a maior frequência de dados ocorreu na faixa de 8,68% a 11,28%, enquanto o pH teve uma frequência elevada em uma faixa estreita de valores entre 7,03 e 7,42, com uma ocorrência significativa também na faixa de 7,94 a 8. Diferentemente do conjunto PAB refinado, os valores de HRT para PAM refinado se concentraram em uma duração de até quase 2 meses. A faixa de OLR com maior frequência foi 3,15-3,91 gVS/(l.d), mas valores tão elevados quanto 11,53 gVS/(l.d) foram encontrados. A distribuição da temperatura foi predominantemente bimodal com a maioria dos dados na temperatura mesófila (cerca de 36°C) e outros próximos dos 55°C (temperatura termofílica). Reatores com volume de aproximadamente 17L foram os mais frequentes, confirmando que os dados foram coletados em escala laboratorial ou piloto. A PAM indicou que o refinamento do conjunto de dados não alterou significativamente o histograma, pois os valores mais baixos ainda foram os mais abundantes e as classes com valores mais altos ainda foram representadas.

Já para o banco de dados para a concentração volumétrica de metano (%CH₄), um conjunto de 1440 amostras foram obtidas após a análise estatística e os resultados dos histogramas para o banco de dados %CH₄ são mostrados na Figura 9 a seguir.



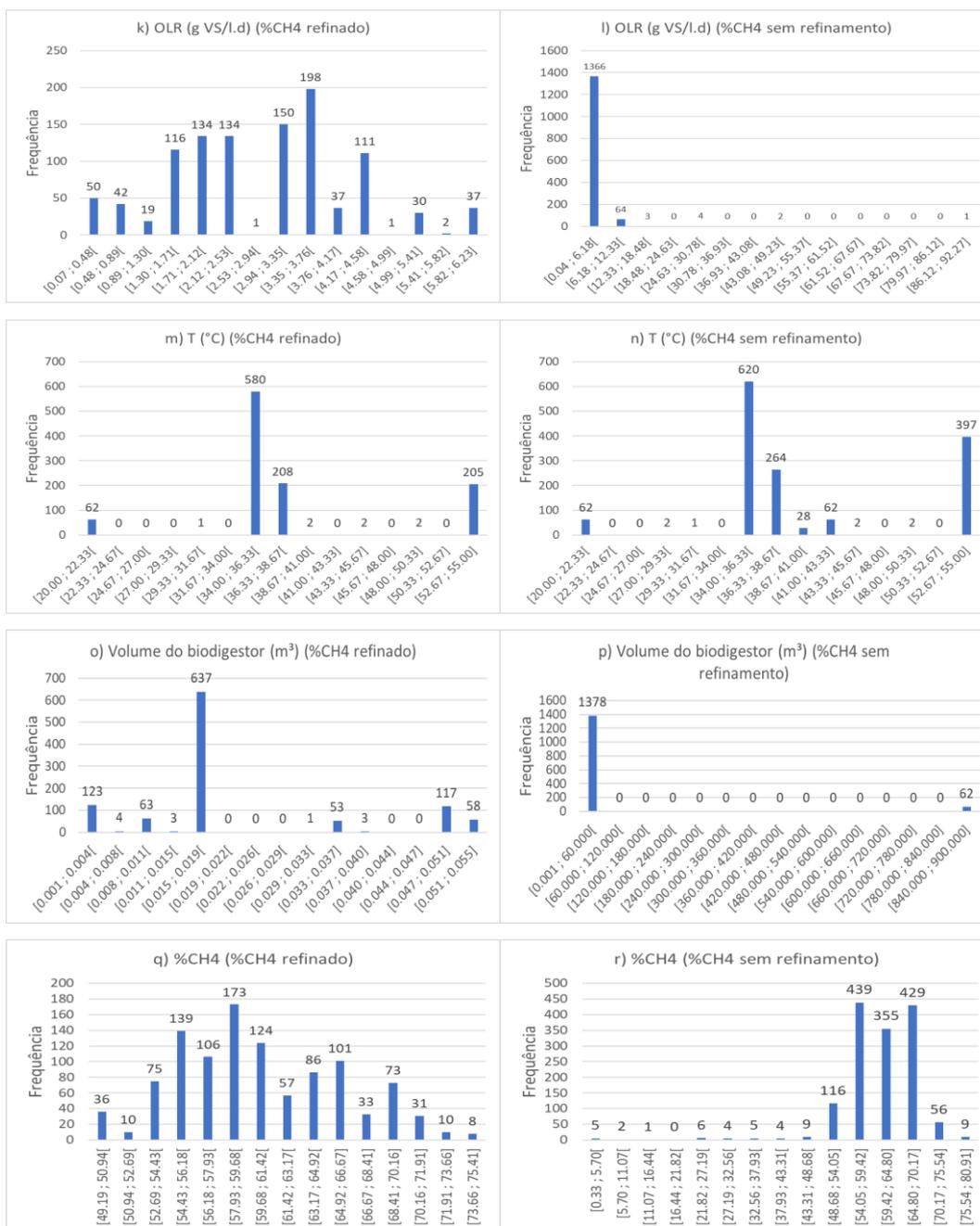


Figura 9 - Histogramas para os bancos de dados %CH₄: comparação das distribuições de “%CH₄ sem refinamento” e “%CH₄ refinado” para tipo de biomassa, tipo de reator, teor de sólidos voláteis (VS), pH, taxa de carga orgânica (OLR), tempo de retenção hidráulica (HRT), temperatura, volume do reator e concentração volumétrica de metano

Durante a análise dos dados %CH₄ sem refinamento, pode-se perceber que algumas informações tinham pouca ou nenhuma representatividade no conjunto de dados. Por exemplo, os tipos de biodigestor semi contínuo de 3 estágios e batelada de 2 estágios, HRT acima de 140 dias, os valores de VS acima de 29,86 %, OLR

acima de 12,33 gVS/l.d e volume de reator acima de 60 m³ foram pouco frequentes nos dados. O histograma de concentração de CH₄ mostrou que os dados com valores abaixo de 48,68% e acima de 75,54% estavam pouco representados no conjunto do banco. O refinamento envolveu a remoção de dados de OLR acima de 6,42 gVS/l.d (valores muito altos foram obtidos para um pseudo-OLR para reator tipo batelada), HRT acima de 91 dias, VS acima de 29,86%, volumes de reator acima de 0,1 m³ e valores de %CH₄ abaixo de 48,68% e acima de 75,54%. Essa etapa gerou o banco %CH₄ refinado com 1062 amostras, 73,75% do conjunto de dados do banco %CH₄ sem refinamento.

Os histogramas do banco de dados %CH₄ refinado mostraram que a codigestão foi pouco representativa, assim como observado para as outros bancos de dados analisados anteriormente. Os tipos de reatores batelada de 2 estágios, semi contínuo de 3 estágios e contínuo de 1 estágio também tiveram pouca representação no banco de dados. Quanto ao teor de VS, os dados coletados estão apresentando uma distribuição quase homogênea na faixa de 3,48% a 9,46%, enquanto o pH segue uma distribuição quase normal centrada na faixa de 7,16-7,29. Assim como para o conjunto PAB refinado, os valores de HRT para %CH₄ refinado se concentraram em uma duração de até 1 mês, conforme relatado na literatura. O histograma de OLR apresenta uma boa dispersão de dados com maior frequência na faixa de 1,3-4,58 gVS/(l.d). A distribuição da temperatura apresenta a maioria dos dados na temperatura mesófila (cerca de 36°C), com dados representativos também na temperatura próxima a 20°C (temperatura psicofrísica) e outros próximos dos 55°C (temperatura termofílica). Reatores com volume de aproximadamente 17L foram os mais frequentes, confirmando que os dados foram coletados em escala laboratorial ou piloto. O histograma de %CH₄ indicou que o refinamento do conjunto de dados alterou significativamente o histograma, pois os valores mais baixos foram cortados possibilitando uma distribuição mais homogênea dos dados.

4.2 Treinamento, validação e teste das redes dos bancos completos

Após a consolidação dos bancos de dados e com a definição das arquiteturas das redes a serem usadas, foi possível iniciar a etapa de treinamento e testes das

mesmas. Para esta etapa, o programa em Matlab desenvolvido testa as diversas topologias individualmente, buscando aquela que apresenta a melhor resposta em termos do coeficiente de determinação R^2 , indicativo do grau de aproximação entre o valor alvo e o valor predito. O algoritmo combina as topologias inseridas para determinar a configuração mais adequada. Ainda, busca-se obter o menor erro médio quadrático (SSE), que juntamente com o R^2 serão as métricas avaliadas. Para as 3 diferentes RNAs estudadas foram 8 neurônios na camada de entrada, que correspondem às variáveis avaliadas no processo de produção de biogás.

4.2.1 Produção acumulada de biogás

Das 1224 amostras presentes no banco PAB refinado, 734 amostras (60%) foram utilizadas para compor o banco de dados utilizado para o treinamento, 245 amostras (20%) foram utilizadas para compor o banco de dados utilizado para validação e outras 245 amostras (20%) usadas para o banco de dados de teste. Definidos os bancos de treinamento, validação e teste, os pesos e as bias de cada topologia foram determinados usando o software Matlab e os resultados são apresentados na Tabela 9.

A Tabela 9 apresenta os resultados do desempenho dos algoritmos Levenberg-Marquardt backpropagation (trainlm), Bayesian regularization backpropagation (trainbr) e Gradient descent (traingdx) na predição de um conjunto de dados de treinamento e teste. Observa-se que o ajuste dos parâmetros usando o algoritmo trainbr resultou na melhor topologia, com boas predições demonstradas pelos valores de SSE e R^2 para o conjunto de treinamento, com valores de R^2 variando entre 0,9926 e 0,9950. Além disso, esse algoritmo apresentou um desempenho semelhante para os valores de R^2 do conjunto de dados de teste, que variaram de 0,9879 a 0,9938. Já o algoritmo trainlm apresentou um desempenho similar ao trainbr, com valores de R^2 do treinamento variando de 0,9914 a 0,9969 e valores de R^2 do teste entre 0,9659 e 0,9935. Por outro lado, o algoritmo traingdx apresentou um desempenho abaixo dos demais algoritmos, com valores similares para os valores de R^2 do treinamento e do teste, variando de 0,6490 a 0,9346 para o treinamento e de 0,6428 a 0,9151 para o teste. No que se refere às funções de ativação e ao número total de neurônios da camada oculta, não foi possível

identificar um comportamento característico que permitisse destacar um valor otimizado para esses parâmetros.

Tabela 10 - Topologias das RNAs para produção acumulada de biogás. Fonte: Elaboração própria.

Identificação da RNA	Algoritmo de treinamento	Função de ativação oculta	Total de neurônios	SSE	R2 Treino	R2 Teste
net1	traingdx	tansig	8	2,390839	0,934586	0,915081
net2			9	3,552616	0,902656	0,875451
net3			10	3,289597	0,909865	0,902455
net4			11	4,760034	0,869711	0,863290
net5		logsig	8	5,868748	0,838885	0,802050
net6			9	12,77789	0,649012	0,642873
net7			10	8,432077	0,768584	0,724202
net8			11	9,880544	0,729334	0,685628
net9	trainbr	tansig	8	0,232253	0,993620	0,991890
net10			9	0,221845	0,993906	0,989483
net11			10	0,204561	0,994381	0,991431
net12			11	0,180821	0,995033	0,993846
net13		logsig	8	0,267915	0,992641	0,987895
net14			9	0,242468	0,993340	0,990066
net15			10	0,213230	0,994143	0,991956
net16			11	0,196803	0,994594	0,993508
net17	trainlm	tansig	8	0,313373	0,991392	0,977798
net18			9	0,230847	0,993659	0,988436
net19			10	0,229751	0,993689	0,991666
net20			11	0,112832	0,996901	0,965889
net21		logsig	8	0,276512	0,992404	0,989782
net22			9	0,232311	0,993619	0,989048
net23			10	0,223351	0,993865	0,979290
net24			11	0,170792	0,995308	0,993541

Para selecionar a melhor performance de rede neural, o critério definido foi escolher a rede que apresentou uma combinação de valores máximos para o R^2 do treinamento e do teste e a que apresentou valor mínimo de *SSE*. Desse modo, a RNA identificada como net12 (*trainbr_tansig_11_purelin_1*) foi selecionada como a melhor performance de rede neural para o banco de dados PAB por conter os melhores valores (R^2 mais próximo a 1). O valor de *SSE* de 0,1808 é considerado um *SSE* baixo, indicando baixo grau de variabilidade dentro do conjunto de dados.

A Figura 10 exibe gráficos comparativos das saídas ("*outputs*") versus os valores-alvo ("*targets*"), bem como das saídas reais e preditas para cada amostra de

dados ao utilizar a rede neural *net12* nos conjuntos de treinamento, validação e teste. Nos gráficos superiores da Figura 10, o valor alvo ou "*target* T" representa a produção acumulada de biogás no banco de dados, enquanto a saída ("*Output* Y") é a previsão feita pela RNA para esse alvo T. O ajuste linear é uma regressão entre a saída ("*Output* Y") e o valor alvo ("*Target* T"), sendo idealmente representado por $Y = T$.

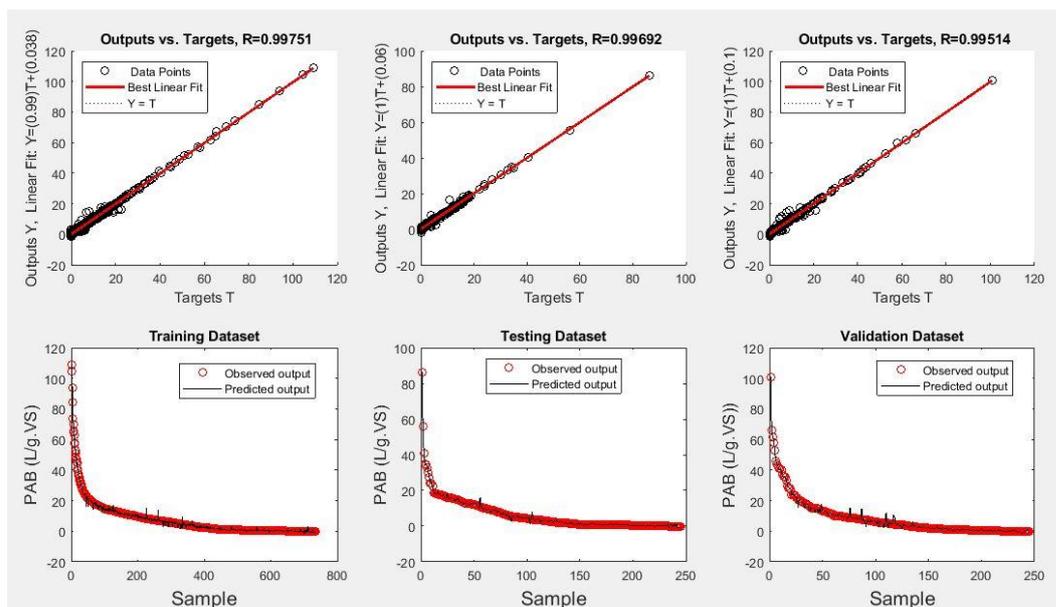


Figura 10 - Resultados para treinamento, teste e validação da RNA. Fonte: Elaboração própria.

Para o treinamento e teste, é possível observar que o gráfico de *outputs vs targets* apresenta uma reta próxima de $Y=X$, indicando assim uma capacidade de previsão boa. O mesmo pode ser observado nos gráficos de *training*, *testing* e *validation dataset*, onde a saída prevista é próxima da saída observada para a maioria das amostras. Quando aplicado ao conjunto de dados da validação, o modelo continuou apresentando um desempenho similar ao treinamento e teste, indicando assim um resultado favorável levando em conta a variedade do banco de dados. É possível considerar que o modelo é capaz de prever bem o volume acumulado de biogás produzido para os insumos considerados, obtendo um R^2 igual a 0,9938 ao aplicar a RNA no banco de dados completo. Existem vários estudos na literatura que utilizam o coeficiente de determinação (R^2) para avaliar o desempenho de redes neurais artificiais (RNAs) na previsão da produção de biogás. Por exemplo, Tian et al. (2021) avaliou o desempenho de RNAs utilizando o R^2 em

um conjunto de dados de diferentes substratos para produção de biogás, obtendo valores de R^2 superiores a 0,9 em alguns casos (Tian et al., 2021). Em outro estudo de 2018, os autores também utilizaram o R^2 para avaliar a capacidade de previsão de uma RNA em relação à produção de biogás a partir de diferentes substratos, obtendo resultados entre 0,8 e 0,99 (Huang et al., 2018). Desse modo, os resultados obtidos neste estudo são consistentes com estudos anteriores sobre o mesmo tema, o que confere maior validade aos achados. Os valores apresentados são robustos e foram obtidos por meio de uma metodologia rigorosa, que levou em conta as limitações e variáveis relevantes para o objeto de estudo. Dessa forma, podemos afirmar que os resultados obtidos são confiáveis e contribuem para o avanço do conhecimento na área.

4.2.2 Produção acumulada de metano

Das 990 amostras presentes no banco de PAM refinado, 594 amostras (60%) foram utilizadas para compor o banco de dados utilizado para o treinamento, 198 amostras (20%) foram utilizadas para compor o banco de dados utilizado para validação e outras 198 amostras (20%) usadas para o banco de dados de teste. Após a definição dos bancos de treinamento, teste e validação, a rede neural foi implementada e a Tabela 10 exibe o resultado de cada topologia utilizada.

Na Tabela 10, pode ser observado que o algoritmo *Levenberg-Marquardt backpropagation* (trainlm) resultou na melhor topologia, com boas predições demonstradas pelos valores de SSE e R^2 para o conjunto de treinamento, com valores de R^2 variando entre 0,9958 e 0,9963. Para o conjunto de dados de teste, esse algoritmo apresentou um decaimento nos valores de R^2 , variando de 0,0050 a 0,9844. Para o algoritmo *Bayesian regularization backpropagation* (trainbr), os valores de R^2 do treinamento variaram de 0,9950 a 0,9959, apresentando um desempenho similar ao algoritmo trainlm. Para o R^2 do teste, os valores ficaram entre 0,8665 e 0,9859. O algoritmo *Gradient descent* (traingdx) apresentou desempenho abaixo dos demais algoritmos, com valores similares para os valores de R^2 do treinamento e do teste, com intervalos que variam de 0,7498 a 0,9649 para o treinamento e 0,7079 a 0,9555 para o teste. Com relação as funções de ativação e

do número total de neurônio da camada oculta não é possível obter diferentes análises.

Tabela 11 - Topologias das RNAs para produção acumulada de metano. Fonte: Elaboração própria

Identificação da RNA	Algoritmo de treinamento	Função de ativação oculta	Total de neurônios	SSE	R2 Treino	R2 Teste
net1	trainidx	tansig	8	1,426920	0,945188	0,928171
net2			9	1,747255	0,932871	0,898692
net3			10	1,149123	0,955813	0,922163
net4			11	0,912067	0,964954	0,955501
net5		logsig	8	6,585548	0,749811	0,707945
net6			9	3,943830	0,848975	0,824389
net7			10	6,284560	0,760118	0,753123
net8			11	4,215128	0,840716	0,742752
net9	trainbr	tansig	8	0,128486	0,995056	0,975380
net10			9	0,117713	0,995470	0,984770
net11			10	0,118037	0,995458	0,875049
net12			11	0,105459	0,995942	0,980389
net13		logsig	8	0,108941	0,995808	0,985874
net14			9	0,114512	0,995594	0,894723
net15			10	0,121063	0,995342	0,866487
net16			11	0,106108	0,995917	0,899939
net17	trainlm	tansig	8	0,109707	0,995779	0,872403
net18			9	0,095649	0,996319	0,793083
net19			10	0,117434	0,995481	0,978568
net20			11	0,114303	0,995602	0,975476
net21		logsig	8	0,136619	0,994743	0,005002
net22			9	0,123773	0,995237	0,861751
net23			10	0,117366	0,995484	0,984456
net24			11	0,098952	0,996192	0,979892

Considerando a combinação de valores máximos para o R^2 do treinamento e do teste e o valor mínimo de *SSE*, a RNA identificada como net13 (*trainbr_logsig_8_purelin_1*) foi selecionada como a melhor performance de rede neural para o banco de dados PAM refinado por apresentar melhores valores de R^2 . O valor de *SSE* de 0,1089 pode ser considerado um valor baixo indicando que a rede está se aproximando mais dos valores observados dentro do conjunto de dados.

Os gráficos Q-Q são apresentados na Figura 11, além das saídas observadas e previstas para cada amostra ao executar a melhor performance de rede neural net13 para o conjunto de dados de treinamento, validação e teste.

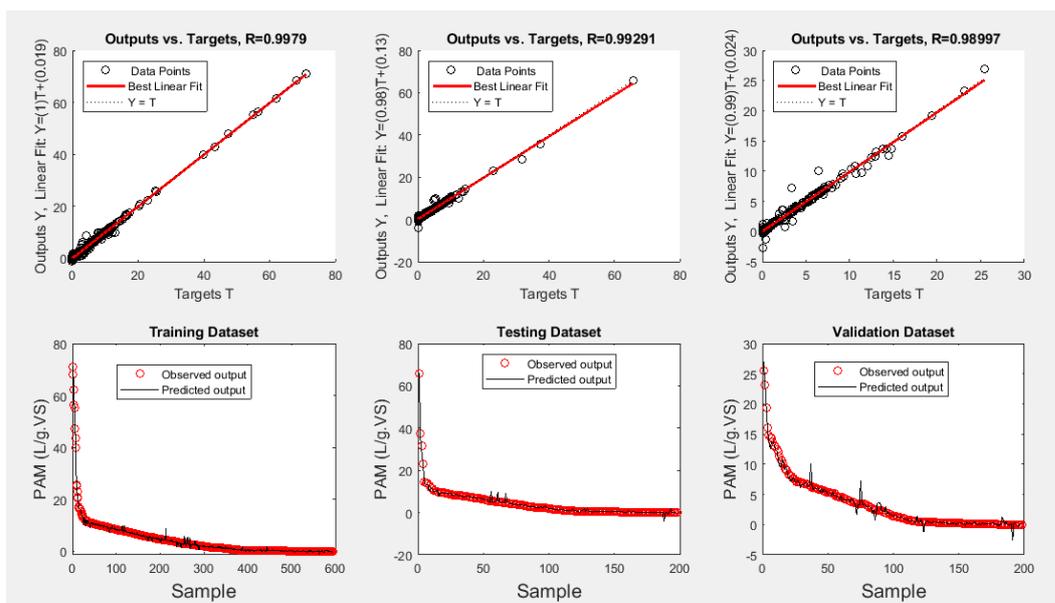


Figura 11 Resultado do treinamento, teste e validação da RNA. Fonte: Elaboração própria.

Na Figura 11, é possível observar que para o treinamento e teste o gráfico de *outputs vs targets* apresenta uma reta próxima de $Y=X$, indicando assim uma capacidade de previsão boa, com um desempenho muito próximo de 1 para treino, teste e validação. O mesmo pode ser observado nos gráficos de *training, testing e validation dataset*, onde a saída prevista é geralmente próxima da saída observada para a maioria das amostras. Quando aplicado ao conjunto de dados da validação, o modelo continuou apresentando um desempenho similar ao treinamento e teste, indicando assim um resultado favorável levando em conta a variedade do banco de dados. É possível considerar que o modelo é capaz de prever bem o volume produzido de metano para os insumos considerados, obtendo um R^2 igual a 0,9926 ao aplicar a RNA no banco de dados completo.

O R^2 e o SSE são amplamente utilizados como métricas de avaliação para RNAs em diferentes áreas, incluindo a produção de biogás e metano. Para a produção de metano, vários estudos têm utilizado o R^2 como métrica de avaliação do desempenho de RNAs. Por exemplo, em Zhang et al. (2020), os autores relataram um valor de R^2 de 0,95 para um modelo de RNA aplicado a dados de produção de metano a partir de resíduos alimentares. Em outro estudo, Tchakouté et al. (2018), os autores obtiveram um valor de R^2 de 0,99 para um modelo de RNA aplicado a dados de produção de metano a partir de lodo de esgoto. Já o SSE não é tão frequentemente utilizado como métrica de avaliação para RNAs na produção de

biogás, mas pode ser útil em alguns casos específicos. É importante lembrar que a escolha da métrica depende da natureza do problema em questão e do objetivo do estudo. Portanto, a comparação dos resultados obtidos nesse presente estudo com os resultados existentes na literatura para desempenho de RNA na produção de metano dependerá da disponibilidade de estudos semelhantes que utilizaram as mesmas métricas de avaliação. No entanto, é possível afirmar que um valor de R^2 igual a 0,9963 indica um desempenho bastante elevado do modelo de RNA na previsão da produção de metano a partir dos insumos considerados nesse trabalho. A presente rede neural apresentou um resultado similar se comparada ao desempenho da rede neural para produção acumulada de biogás, visto anteriormente.

4.2.3 Concentração volumétrica de metano (CH₄) no biogás

No conjunto de dados de concentração de CH₄ refinado utilizado para esta análise, havia um total de 1062 amostras disponíveis. Destas, 60% (638 amostras) foram selecionadas para compor o conjunto de dados de treinamento, enquanto 20% (212 amostras) foram reservadas para o conjunto de dados de validação e outras 20% (212 amostras) foram utilizadas para compor o conjunto de dados de teste. Com os conjuntos de dados definidos, a RNA foi implementada utilizando o software Matlab e os resultados de cada topologia utilizada estão apresentados na Tabela 11.

Da Tabela 11, pode ser observado que o algoritmo *Bayesian regularization backpropagation* (trainbr) apresentou a melhor topologia, com boas previsões demonstradas pelos valores de SSE e R^2 para o conjunto de treinamento, com valores de R^2 variando entre 0,8993 e 0,9133. Para o conjunto de dados de teste, esse algoritmo apresentou um considerável decaimento nos valores de R^2 , variando de 0,6470 a 0,8460. Para o algoritmo *Levenberg-Marquardt backpropagation* (trainlm) os valores de R^2 do treinamento variaram de 0,7498 a 0,9250, apresentando uma variação maior e desempenho abaixo do algoritmo trainbr. Para O R^2 do teste os valores ficaram entre 0,0331 e 0,8244. O algoritmo *Gradient descent* (traingdx) apresentou desempenho abaixo dos demais algoritmos, com valores similares para os valores de R^2 do treinamento e do teste, com intervalos

que variam de 0,7416 a 0,8379 para o treinamento e 0,6405 a 0,8135 para o teste. Não foi observado um comportamento característico que pudesse indicar um valor otimizado para os parâmetros relacionados às funções de ativação e ao número total de neurônios na camada oculta.

Tabela 12 - Topologias das RNAs para concentração volumétrica de CH₄ no biogás. Fonte: Elaboração própria

Identificação da RNA	Algoritmo de treinamento	Função de ativação oculta	Total de neurônios	SSE	R ² Treino	R ² Teste
net1	trainidx	tansig	8	18,78124	0,837918	0,813536
net2			9	24,55549	0,788046	0,712045
net3			10	25,43835	0,780408	0,757098
net4			11	21,38735	0,815430	0,756993
net5		logsig	8	25,42367	0,780572	0,720175
net6			9	24,89978	0,785147	0,736515
net7			10	28,22501	0,756373	0,705690
net8			11	29,93670	0,741604	0,640566
net9	trainbr	tansig	8	11,66587	0,899307	0,806417
net10			9	11,09730	0,904215	0,749590
net11			10	11,25516	0,902864	0,727512
net12			11	10,03863	0,913361	0,750583
net13		logsig	8	11,66006	0,899366	0,751286
net14			9	11,64576	0,899489	0,795372
net15			10	11,45034	0,901177	0,846046
net16			11	10,66938	0,907923	0,647048
net17	trainlm	tansig	8	11,77984	0,898312	0,746665
net18			9	12,00977	0,896327	0,473228
net19			10	9,965813	0,913971	0,686480
net20			11	28,97860	0,749846	0,683656
net21		logsig	8	11,92015	0,897101	0,824454
net22			9	10,68182	0,907791	0,802735
net23			10	8,945435	0,922780	0,033120
net24			11	8,680672	0,925065	0,434035

Com o intuito de selecionar a rede neural com melhor desempenho, foi adotado um critério que consistiu em escolher a rede que apresentasse a combinação máxima de valores para o R² do treinamento e do teste, e o valor mínimo de SSE. Como resultado, a rede neural identificada como net15 (*trainbr_logsig_10_purelin_1*) foi escolhida como a que teve melhor performance para o banco de dados de concentração de CH₄, devido aos seus valores de R² mais

próximos de 1. Com base no valor de SSE de 11,4503, pode-se inferir que os dados do conjunto apresentam uma baixa variabilidade.

A RNA concentração volumétrica de CH_4 é a que apresenta performance inferior se comparada com as demais RNA aqui estudadas, resultado que pode estar relacionado a caracteriza de construção do banco de dados que apresenta saída com dados com pouco variação. Desse modo, foi possível observar durante os testes que a RNA sofreu maior influência dos valores outliers das variáveis, o que prejudicou seu desempenho. O valor de SSE para essa rede foi maior se comparado aos valores de SSE das redes neurais para produção acumulada de biogás e para a produção acumulada de metano visto anteriormente. Isso mostra um maior grau de variabilidade dentro do conjunto de dados para concentração volumétrica de metano.

A Figura 12 mostra os gráficos das previsões versus o alvo, bem como das previsões e do alvo real para cada amostra de dado ao aplicar a rede neural net15 para os dados de treinamento, teste e validação. Nos gráficos superiores da Figura 17, o alvo ou target T representa a produção acumulada de biogás no banco de dados, enquanto a previsão (“Output Y”) é a estimativa feita pela RNA do alvo T. O ajuste linear é a regressão linear entre Output Y e Target T e sua forma ideal seria $Y=T$.

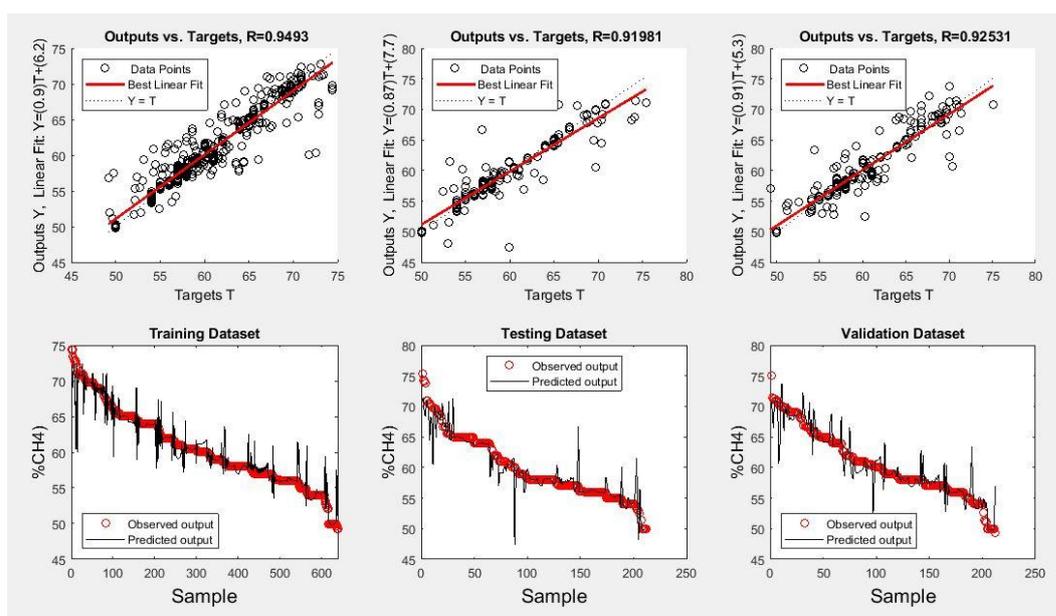


Figura 12 - Resultado do treinamento, teste e validação da RNA para CH_4 . Fonte: Elaboração própria

De acordo com a Figura 12, é possível observar que para o treinamento e teste o gráfico de *outputs vs targets* apresenta uma reta próxima de $Y=X$, indicando assim uma capacidade de previsão boa, com um desempenho menor alcançado para o teste. O mesmo pode ser observado nos gráficos de *training, testing e validation dataset*, onde a saída prevista é geralmente próxima da saída observada para a maioria das amostras. Quando aplicado ao conjunto de dados da validação, o modelo continuou apresentando uma performance boa, similar ao treinamento e teste, indicando assim um resultado favorável levando em conta o tamanho grande do banco de dados. É possível considerar que o modelo é capaz de prever a concentração volumétrica de CH_4 de forma satisfatória para os insumos considerados, obtendo um R^2 igual a 0,8809 ao aplicar a RNA no banco de dados completo.

A utilização de modelos de rede neural artificial para previsão tem sido amplamente investigada e estudada. Estudos anteriores, como o realizado por Gonçalves Neto et al. (2020), obtiveram resultados de R^2 de 0,993, 0,848 e 0,866 para treino, teste e validação, respectivamente, para topologias semelhantes as utilizadas no presente trabalho. Em outro estudo, realizado por Mougari et al. (2021), obteve-se um desempenho de treino, teste e validação igual a 0,999 para uma topologia de rede neural composta por 4 camadas de entradas, 2 camadas ocultas e uma camada de saída para a produção de biogás. Diante disso, os resultados de desempenho de redes neurais encontrados neste presente estudo estão em consonância com os resultados apresentados na literatura para casos similares. Esses achados destacam a eficácia do uso de modelos de rede neural artificial como uma ferramenta de previsão para análise de dados em diversas áreas de pesquisa.

4.3 Superfícies de resposta para bancos de dados completos

4.3.1 Superfícies de resposta para a produção acumulada de biogás

Foi escolhido um valor para as variáveis contínuas invariantes em cada simulação de superfície, o qual correspondeu à média da categoria de maior representatividade no banco de dados. A variável tipo de reator/alimentação e tipo de biomassa foram fixadas de acordo com o caso de maior ocorrência

respectivamente analisados. É válido ressaltar que os valores de pH e temperatura utilizados correspondem a condições de reação mesofílica e o HRT foi fixado em 30 dias, que é um valor comum na literatura para a produção de biogás a partir de resíduos alimentares (Beux, 2005). Valores mostrados na Tabela 12.

Tabela 13 - Valores fixos das variáveis. Fonte: Elaboração própria

Variável de entrada	Valor fixo
VS (%)	8,700
pH	7,200
OLR (g VS/ (l.d))	3,170
HRT (d)	30,00
Temperatura (°C)	35,00
Volume do reator (m³)	0,018
Tipo de Reator/alimentação	-1
Tipo de biomassa	-1

Para manter a simulação das superfícies de acordo com a realidade do banco de dados PCB, foram estabelecidos valores mínimos e máximos para cada variável de entrada. O valor mínimo corresponde ao menor valor registrado da variável no banco de dados PAB para o treinamento, enquanto o valor máximo corresponde ao maior valor registrado da variável de entrada no mesmo banco de dados. A Tabela 13 apresenta esses valores e também os valores mínimos e máximos para a saída PAB.

Tabela 14 - Valores de máximo e mínimo das variáveis para o banco de dados PAB treinamento.

Fonte: Elaboração própria

Variável de entrada	Mínimo	Máximo
Tipo de biomassa	-1	1
Reator/alimentação	-3	3
VS (%)	1,330	25,91
pH	6,400	8,000
OLR (g VS/ (l.d))	0,1900	8,700
HRT (d)	0,000	92,81
Temperatura (°C)	20,000	55,000
Volume do biodigestor (m³)	0,0004	0,0640
Variável de saída		
Produção acumulada de biogás (L/g VS)	0	109,17

Da análise dos efeitos de cada variável, foi possível perceber que alguns efeitos combinados podem influenciar o sistema, desse modo, foram geradas superfícies de resposta para essas variáveis combinadas e são mostradas nas Figuras 13 a 21 a seguir. As Figuras 13, 14 e 15 acima mostram as superfícies de resposta para a combinação das variáveis pH e volume de reator.

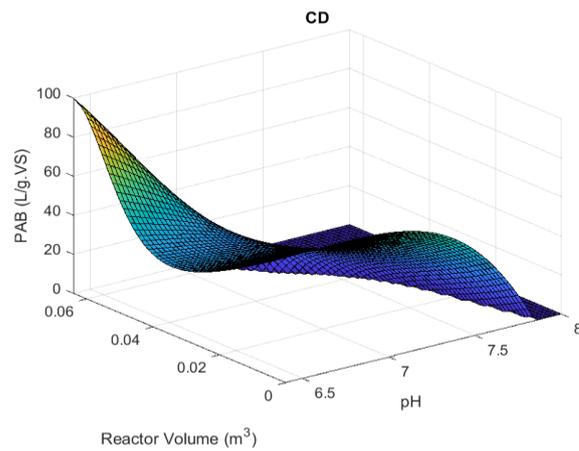


Figura 13 - Superfície de resposta para a combinação entre pH e volume de reator para CD

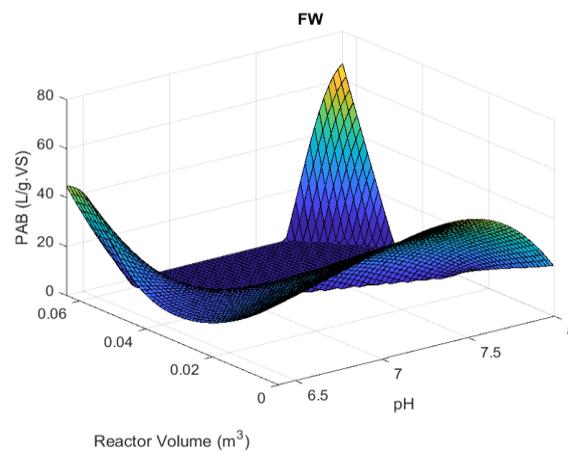


Figura 14 - Superfície de resposta para a combinação entre pH e volume de reator para FW

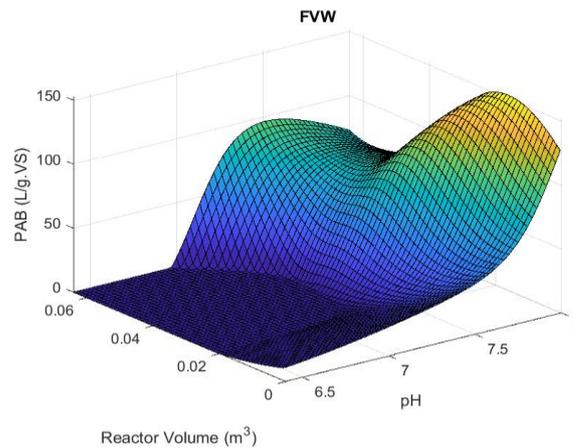


Figura 15 - Superfície de resposta para a combinação entre pH e volume de reator para FVW

Para o substrato FVW é possível observar que a máxima produção acumulada de biogás é atingida para valores de pH termofílicos, em torno de 8. A RNA treinada apresentou resultados consistentes para FW e CD para volumes até 0,020 m³ (ou seja, a faixa com mais dados disponíveis). Para valores maiores, os resultados obtidos não correspondem ao esperado para um biodigestor, embora alguns dados estejam nessa região. Para FVW, é preciso fazer uma observação adicional, pois os valores estimados ultrapassaram o valor máximo de PAB nos dados para os menores valores de volume. Portanto, neste momento, o modelo de RNA é mais confiável na região com volume menor que 20 L para FW e CD. Na literatura científica, diversos estudos já foram realizados para avaliar a influência do pH na produção de biogás em diferentes tipos de reatores e substratos. De acordo com Deressa et al. (2015) a digestão anaeróbica é muito sensível à mudança no pH e é importante manter o pH de 6,7-7,4 para um sistema saudável. Portanto, os resultados encontrados nesta pesquisa estão de acordo com a literatura e reforçam a importância do controle do pH para maximizar a produção de biogás em sistemas anaeróbios. As inconsistências encontradas aqui serão melhor exploradas na seção 4.4 do documento.

As Figuras 16, 17 e 18 abaixo apresentam as superfícies de resposta para a interação entre a temperatura e o volume do reator.

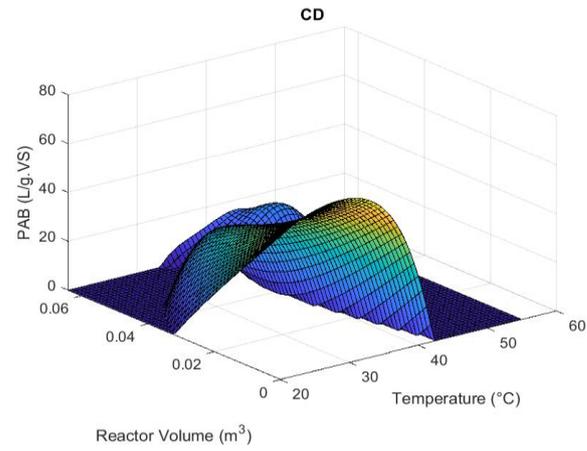


Figura 16 - Superfície de resposta para a combinação entre temperatura e volume de reator para CD

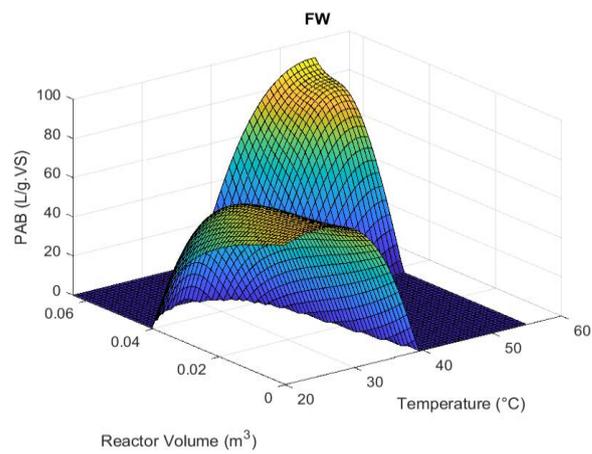


Figura 17 - Superfície de resposta para a combinação entre temperatura e volume de reator para FW

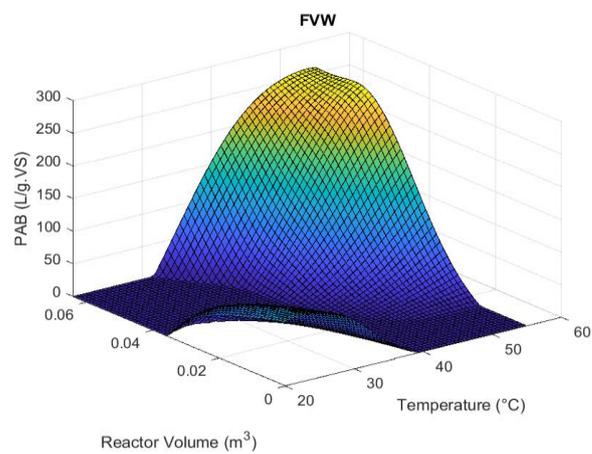


Figura 18 - Superfície de resposta para a combinação entre temperatura e volume de reator para FVW

As superfícies para FW e CD mostraram valores coerentes de PAB e indicaram condições ótimas em temperatura mesófila para volume baixo e temperatura termofílica para volume maior que $0,04 \text{ m}^3$, respectivamente. Para FW, os valores estimados superaram o valor máximo de PAB no banco de dados para os maiores valores de volume e temperatura, indicando a necessidade um melhor tratamento dos dados, que será feito na seção seguinte. Os estudos mostram que a temperatura ótima para produção de biogás varia de 28°C a 37°C (Adamma et al., 2023; Rattanapan et al., 2019). Portanto, pode-se concluir que a temperatura desempenha um papel crucial na produção de biogás a partir de diferentes substratos.

As Figura 19, 20 e 21 mostram a variação da PAB em L/gVS em função do VS e do volume do reator:

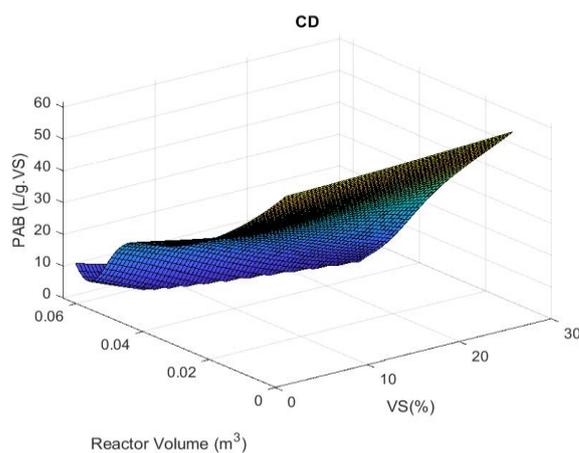


Figura 19 - Superfície de resposta para a combinação entre VS e volume de reator para CD

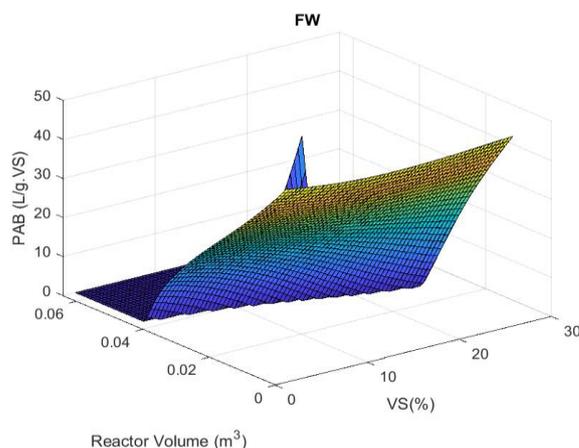


Figura 20 - Superfície de resposta para a combinação entre VS e volume de reator para FW

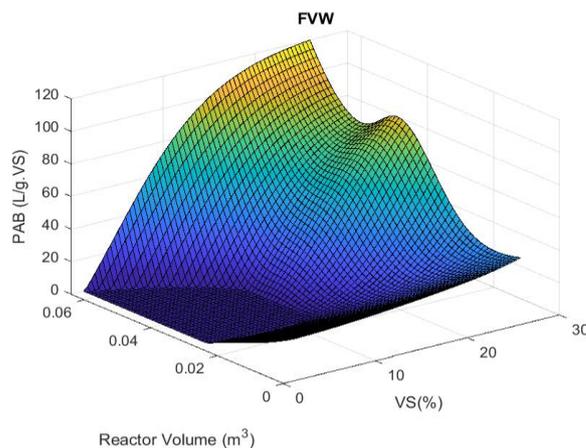


Figura 21 - Superfície de resposta para a combinação entre VS e volume de reator para FVW

Para os substratos FW e CD a produção de biogás é aumentada para pequenos volumes, abaixo de $0,025 \text{ m}^3$, variando muito pouco com o valor de VS. Para FV o pico de produção acumulada de biogás atinge um volume de 40 L/g.VS , enquanto que para CD o pico de PAB é atingido em valores por volta de 50 L/g.VS . É importante observar que para FVW existe uma mudança de comportamento, onde a maior PAB é atingida para valores de volume de reator maiores, acima de $0,025 \text{ m}^3$, e existe uma sensibilidade maior a variação de VS, com valores de PAB aumentando a mediada que o VS aumenta. De acordo com estudos, sólidos voláteis de substratos de resíduos alimentares (FW) e resíduos de frutas e vegetais (FVW) têm uma influência significativa na produção de biogás (Khadka et al., 2022; Hernández-Shek et al., 2016). A codigestão de FW com outros substratos pode melhorar a estabilidade e o potencial do biogás, embora o rendimento total de biogás e o teor de metano possam diminuir com o aumento do teor de substrato de codigestão (Andrés Donoso-Bravo et al., 2021). Portanto, os resultados encontrados na presente etapa do estudo são pouco conclusivos ao não apontar a esperada influência do VS na produção de biogás, justificando assim uma análise dos dados mais específicos, que será feita na seção 4.4 do estudo.

4.3.2 Superfícies de resposta para produção de metano

Assim como para as superfícies de resposta para produção de biogás, para cada simulação de superfície foi escolhido um valor para as variáveis contínuas

invariante o qual correspondeu à média da categoria de maior representatividade no banco de dados de treinamento. A variável tipo de reator/alimentação e tipo de biomassa foram fixadas de acordo com o caso de maior ocorrência respectivamente analisados. É importante notar aqui que os valores de pH e temperatura apontam para condições de reação mesofílica e o valor de HRT fixado em 30 dias. Uma tabela contendo os valores fixados para as variáveis é apresentada na Tabela 14 a seguir:

Tabela 15 - Valores fixos das variáveis. Fonte: Elaboração própria

Variável de entrada	Valor fixo
VS (%)	9,330
pH	7,190
OLR (g VS/ (l.d))	3,180
HRT (d)	30,00
Temperatura (°C)	35,66
Volume do reator (m³)	0,017
Tipo de Reator/alimentação	-1
Tipo de biomassa	-1

Para manter a simulação das superfícies de acordo com a realidade do banco de dados PAM, foram estabelecidos valores mínimos e máximos para cada variável de entrada. O valor mínimo corresponde ao menor valor registrado da variável no banco de dados PAM para o treinamento, enquanto o valor máximo corresponde ao maior valor registrado da variável de entrada no mesmo banco de dados. A Tabela 15 apresenta esses valores e também os valores mínimos e máximos para a saída PAM.

Tabela 16 - Valores de máximo e mínimo das variáveis para o banco de dados PAM treinamento.

Fonte: Elaboração própria

Variável de entrada	Mínimo	Máximo
Tipo de biomassa	-1	1
Reator/alimentação	-3	3
VS (%)	3,480	22,99
pH	7,03	8
OLR (g VS/ (l.d))	0,148	10,25
HRT (d)	0,000	51,59
Temperatura (°C)	35,00	55,00
Volume do biodigestor (m ³)	0,0005	0,0545
Variável de saída		
Produção acumulada de biogás	0	71,11

A partir da análise dos efeitos de cada variável, foi possível perceber que alguns efeitos combinados podem influenciar o sistema, desse modo, foram geradas superfícies de resposta para essas variáveis combinadas e são mostradas nas Figuras 22 a 27 a seguir.

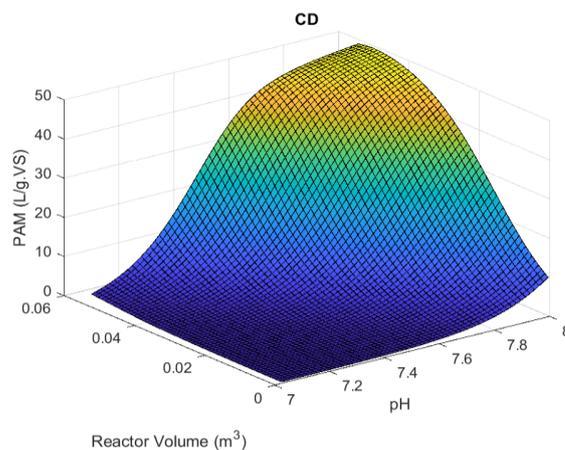


Figura 22 - Superfície de resposta para a combinação entre pH x volume de reator para CD, PAM.

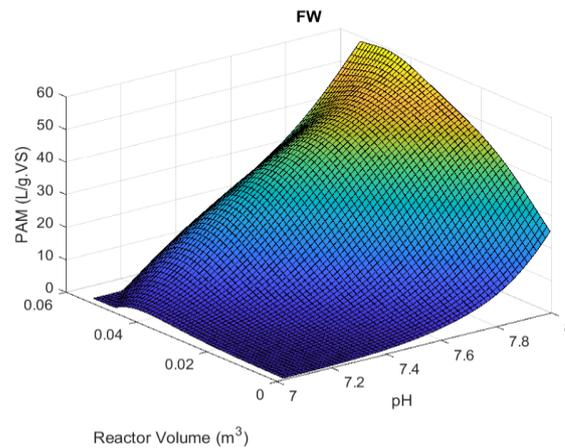


Figura 23 - Superfície de resposta para a combinação entre pH x volume de reator para FW, PAM.

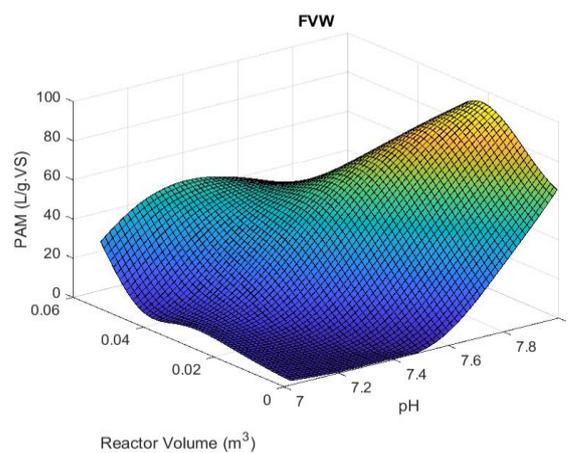


Figura 24 - Superfície de resposta para a combinação entre pH x volume de reator para FVW, PAM.

As Figuras 22, 23 e 24 mostram as superfícies de resposta para a combinação entre o pH e o volume de reator. O pico de produção está presente na região termofílica de pH, valores entre 7,6 de 8, para os três tipos de substratos. Para o tipo de substrato CD e FW, a maior produção acumulada de metano foi atingida para volumes entre 0,04 m³ a 0,06 m³ enquanto que para o tipo de substrato FVW o pico de produção foi atingido para volumes menores de reator, entre 0,02 m³ e 0,04 m³. É possível observar que a influência do tamanho do reator aqui não foi muito determinante para o aumento ou decaimento da produção acumulada de metano uma vez que, de acordo com as superfícies, a variação com essa variável não foi tão marcante, sendo mais afetada pelo valor de pH. Existem diversos estudos na literatura que relatam a influência do pH na produção de metano em

biodigestores. Um estudo realizado por Sosnowski et al. (2018) relatou que o pH ótimo para a produção de metano em biodigestores foi de 7,2. Outro estudo de Liu et al. (2019) também relatou que o pH ótimo para a produção de metano foi de 7,2. Já o estudo de Zhong et al. (2019) encontrou um valor ótimo de pH um pouco mais alto, em torno de 7,5. De acordo com a revisão bibliográfica apresentada, na acetogênese e na metanogênese, o pH ideal fica entre 6,5 à 8,0 (FNR, 2010). Portanto, os resultados obtidos no presente estudo estão de acordo com alguns estudos na literatura e também dentro da faixa ótima para a produção de metano em biodigestores.

As Figuras 25, 26 e 27 abaixo mostram as superfícies de resposta para a combinação entre as variáveis VS e volume de reator.

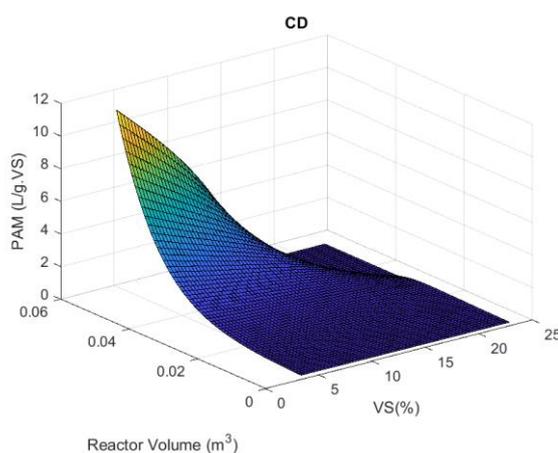


Figura 25 - Superfície de resposta para a combinação entre VS x volume de reator para CD, PAM.

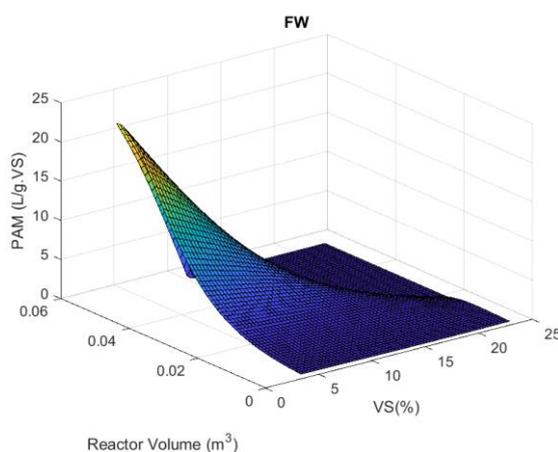


Figura 26 - superfície de resposta para a combinação entre VS x volume de reator para FW, PAM.

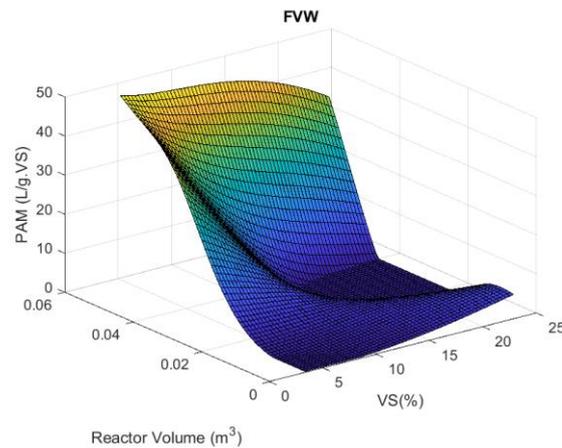


Figura 27 - Superfície de resposta para a combinação entre VS x volume de reator para FVW, PAM.

Nas figuras acima, as superfícies para os 3 tipos de substratos apresentam semelhanças com maior PAM para valores maiores de volume de reator, acima de $0,02 \text{ m}^3$. A região de melhor desempenho para PAM varia de VS entre 2% e 10%, com decaimento de produção para valores acima de 10% para CD e FW e acima de 20% para FVW. Os valores de PAM para o substrato FVW são maiores se comparados com o substrato FW e CD, mas estão de acordo com a faixa de valores presentes no banco de dados. De acordo com a literatura existente, o valor ideal de sólidos voláteis para a produção de metano pode variar dependendo do tipo de substrato utilizado no processo de digestão anaeróbia. Para o substrato food waste (FW), alguns estudos relatam que os valores ótimos de sólidos voláteis podem variar de 5% a 10%, enquanto outros estudos apontam valores mais elevados, em torno de 12% a 15% (Khalid et al., 2011; Mata-Alvarez et al., 2014). Já para o substrato *vegetable food waste* (FVW), há poucos estudos disponíveis na literatura que descrevem valores ideais de sólidos voláteis para produção de metano. No entanto, alguns estudos indicam que valores ótimos de sólidos voláteis podem variar de 8% a 10% (Hu et al., 2019). Portanto, os resultados obtidos no presente estudo, que indicam valores ótimos de sólidos voláteis de 2% a 10% estão de acordo com alguns estudos na literatura que relatam valores similares. No entanto, é importante destacar que outros fatores, como pH, tempo de retenção hidráulica e temperatura, também podem influenciar a produção de metano em biodigestores e devem ser considerados na otimização do processo.

4.3.3 Superfícies de resposta para concentração volumétrica de metano (CH₄) no biogás

Foram investigados os efeitos combinados do tipo de reator/alimentação com pH, pH com OLR e tipo de biomassa com pH. Portanto, a análise dos resultados permite concluir que o processo de biodigestão é influenciado por diversas variáveis individuais e combinações de variáveis, sendo necessário um cuidadoso controle e monitoramento do processo para se obter um alto rendimento de CH₄ na produção de biogás.

Assim como foi feito para as demais respostas, para cada simulação de superfície o valor escolhido para as variáveis contínuas invariantes foi média da categoria de maior representatividade no banco de dados. No caso da variável tipo de reator/alimentação e tipo de biomassa, elas foram fixadas de acordo com o caso de maior ocorrência no banco de dados de treinamento. É importante notar aqui que os valores de pH e temperatura apontam para condições de reação mesofílica e HRT fixo em 30 dias, como nas demais saídas. Uma tabela contendo os valores fixados para as variáveis é apresentada na Tabela 16 a seguir:

Tabela 17 - Valores fixos das variáveis. %CH₄. Fonte: Elaboração própria

Variável de entrada	Valor fixo
VS (%)	8,700
pH	7,200
OLR (g VS/ (l.d))	3,060
HRT (d)	30,00
Temperatura (°C)	35,00
Volume do reator (m ³)	0,018
Tipo de Reator/alimentação	-1
Tipo de biomassa	-1

Foi estabelecido um intervalo de valores mínimo e máximo para cada variável de entrada, com o objetivo de garantir a coerência das superfícies de simulação com o banco de dados. O valor mínimo foi determinado com base no menor valor observado da variável no banco de dados de treinamento, enquanto que o valor máximo foi estipulado de acordo com o maior valor da variável de entrada

observado no mesmo banco de dados. A Tabela 17 apresenta esses valores e também os valores mínimos e máximos para a saída %CH₄.

Tabela 18 - Valores de máximo e mínimo das variáveis para o banco de dados %CH₄ treinamento.

Fonte: Elaboração própria

Variável de entrada	Mínimo	Máximo
Tipo de biomassa	-1	1
Reator/alimentação	-3	3
VS (%)	3,480	25,91
pH	6,400	8,200
OLR (g VS/ (l.d))	0,1900	6,220
HRT (d)	0,000	91,00
Temperatura (°C)	20,00	55,000
Volume do biodigestor (m3)	0,0005	0,0545
Variável de saída		
Produção acumulada de biogás	49,19	75,41

Após avaliar os efeitos individuais de cada variável, foi observado que determinadas combinações desses efeitos podem afetar o sistema. Como resultado, foram criadas superfícies de resposta para essas variáveis combinadas, as quais estão ilustradas nas Figuras abaixo. As Figuras 28, 29 e 30 apresentam os gráficos que mostram como a configuração do reator e o pH influenciam a produção de metano para diferentes tipos de substratos, como FWW, FW e CD.

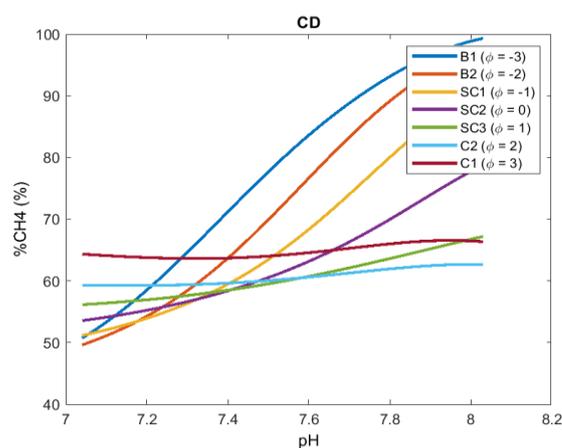


Figura 28 - Superfície de resposta para configuração de reator x pH para CD. B1 (Batelada 1 estágio = -3), B2 (Batelada 2 estágios = -2), SC1 (Semi contínuo 1 estágio = -1), SC2 (Semi contínuo 2 estágios = 0), SC3 (Semi contínuo 3 estágios= 1), C2 (Contínuo 2 estágios = 2), C1 (Contínuo 1 estágio = 3).

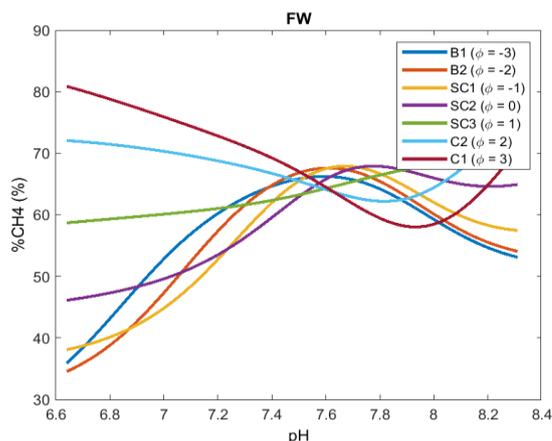


Figura 29 - Superfície de resposta para configuração de reator x pH para FW. B1 (Batelada 1 estágio = -3), B2 (Batelada 2 estágios = -2), SC1 (Semi contínuo 1 estágio = -1), SC2 (Semi contínuo 2 estágios = 0), SC3 (Semi contínuo 3 estágios = 1), C2 (Contínuo 2 estágios = 2), C1 (Contínuo 1 estágio = 3).

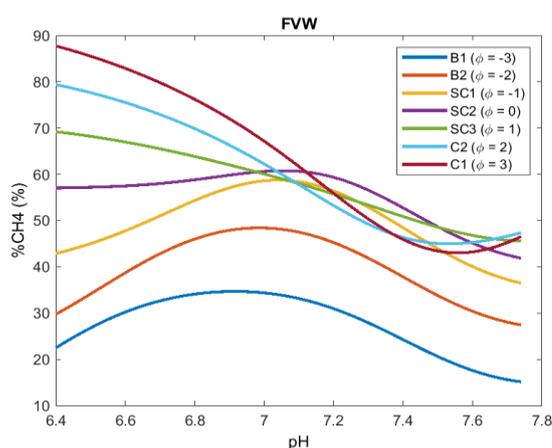


Figura 30 - Superfície de resposta para configuração de reator x pH para FVW. B1 (Batelada 1 estágio = -3), B2 (Batelada 2 estágios = -2), SC1 (Semi contínuo 1 estágio = -1), SC2 (Semi contínuo 2 estágios = 0), SC3 (Semi contínuo 3 estágios = 1), C2 (Contínuo 2 estágios = 2), C1 (Contínuo 1 estágio = 3).

De modo geral, observa-se nas Figuras acima que tanto para FVW quanto para FW, as curvas apresentam um comportamento semelhante, com uma região de pH entre 6,9 e 7,9 apresentando maior concentração de metano para diferentes tipos de reator. Já para CD, há um aumento na concentração de metano na região termofílica de pH segundo a superfície de resposta. Esses resultados são consistentes com a literatura existente, que indica a influência significativa do pH e do tipo de reator na produção de metano a partir de resíduos FW e FVW

(MAGAMA et al., 2022; LOHANI, 2020). De fato, a maioria das bactérias anaeróbicas, incluindo bactérias formadoras de metano, funcionam em uma faixa de pH de 5,5 a 8,5, mas idealmente em um pH de 6,8 a 7,6, e a taxa de produção de metano pode diminuir se o pH for menor que 6,3 ou maior que 7,8 (ABEBE, 2017). No entanto, uma diminuição no pH para uma faixa de 4-5 pode retardar a produção de gás (LOHANI, 2020). Além disso, o tipo de reator também desempenha um papel importante na produção de metano. A codigestão foi considerada a estratégia mais simples de implementar para aumentar o rendimento de metano entre as variáveis como o pH (MAGAMA et al., 2022). Portanto, esses resultados indicam que a otimização da configuração do reator e do pH é crucial para maximizar a produção de metano a partir de resíduos FW e FVW, e a codigestão pode ser uma estratégia promissora para melhorar o rendimento de metano.

As superfícies de resposta apresentadas nas Figuras 31, 32 e 33 mostram a interação entre o pH e o volume de reator na produção de metano a partir dos diferentes tipos de biomassa.

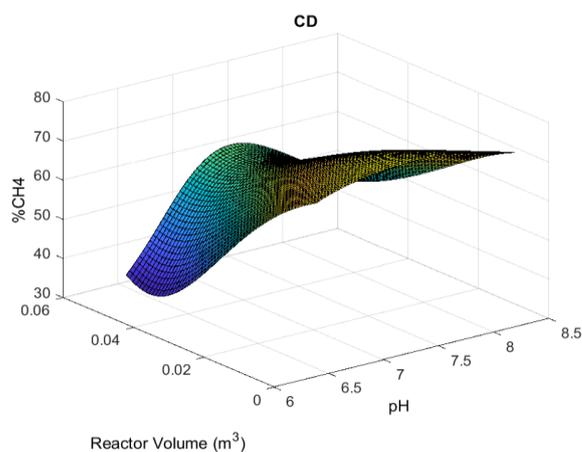


Figura 31 - Superfície de resposta para pH x volume de reator para CD.

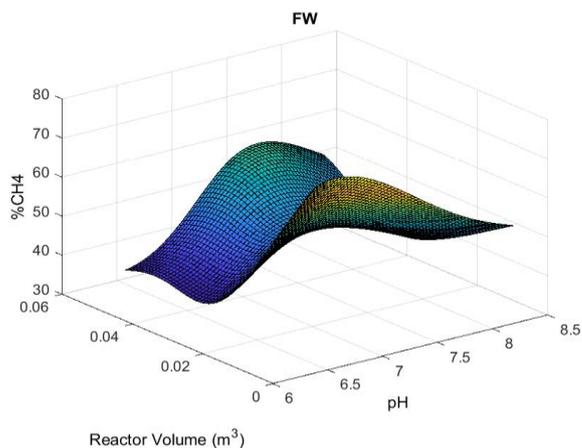


Figura 32 - Superfície de resposta para pH x volume de reator para FW.

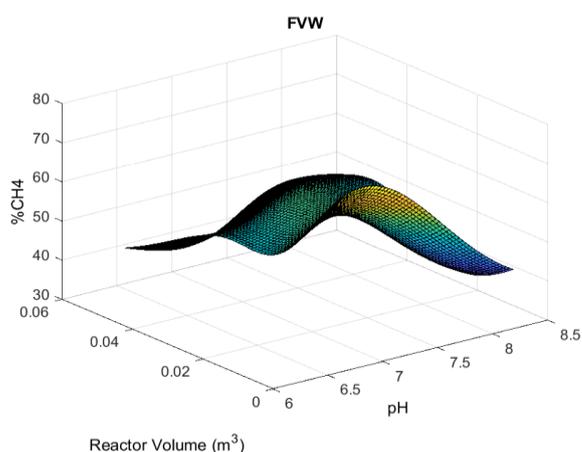


Figura 33 - Superfície de resposta para pH x volume de reator para FVW.

Como observado nas figuras acima, no caso de CD, a superfície de resposta indica que a maior concentração volumétrica de metano é encontrada na fase termofílica de pH e para volumes de reator menores, chegando a um pico de cerca de 70% de CH₄. Esse resultado pode ser explicado pelo fato de que as bactérias termofílicas são mais eficientes na produção de metano a partir de resíduos orgânicos de alta complexidade, como a biomassa de CD (ABDELRAHMAN,2021; ESCAMILLA-ALVARADO, 2010). Já para a biomassa FW, a superfície de resposta mostra picos de concentração de metano para valores de pH em torno de 7. O pico de concentração é em torno de 60% de CH₄. Por fim, para a biomassa FVW, a superfície de resposta apresenta um comportamento parabólico, com maior concentração de metano obtida na faixa de pH em torno de 7 e um pico de concentração por volta de 57% de CH₄. Em resumo, os resultados

das superfícies de resposta são condizentes com os trabalhos da literatura, que apontam uma faixa ótima de pH para a produção de metano entre 6,8 e 7,6 (ABEBE, 2017). Além disso, a interação entre o pH e o volume de reator pode ter um impacto significativo na produção de metano a partir de diferentes tipos de biomassa, sendo necessário avaliar essas variáveis em conjunto para obter uma produção eficiente de biogás.

Nas Figuras 34, 35 e 36 abaixo, é possível observar as superfícies de resposta para a combinação de OLR e volume de reator.

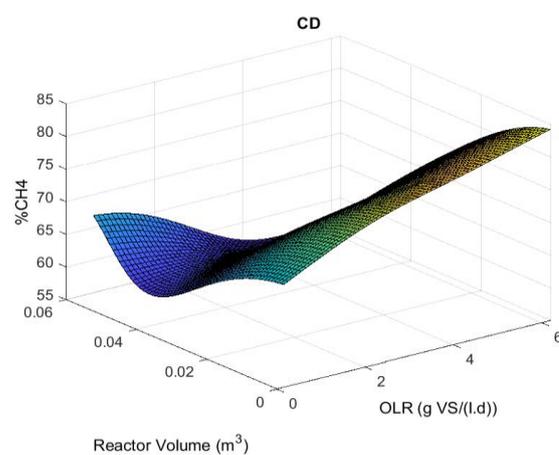


Figura 34 - Superfície de resposta para OLR x volume de reator para CD, %CH4.

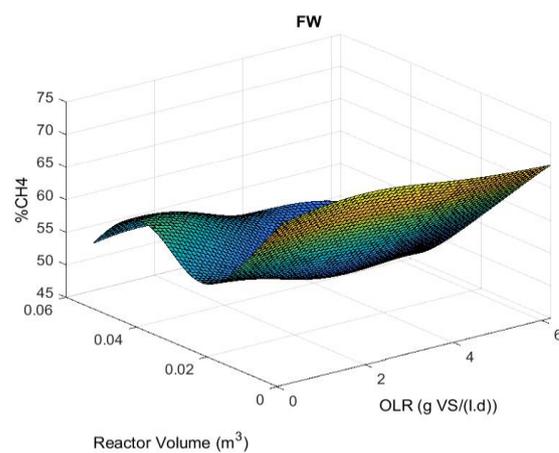


Figura 35 - Superfície de resposta para OLR x volume de reator para FW, %CH4.

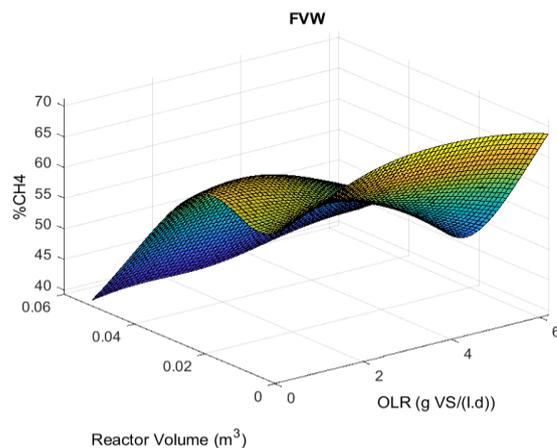


Figura 36 - Superfície de resposta para OLR x volume de reator para FVW, %CH₄.

É possível observar das Figuras acima que para os três tipos de substratos, a maior concentração volumétrica de metano é atingida para volumes de reator menores, abaixo de 0,02 m³. Para o substrato CD, a concentração volumétrica de metano aumenta e atinge um pico à medida o OLR aumenta. Já para o substrato FW, não é verificada uma influência da variação de OLR na resposta, enquanto a superfície de resposta para FVW apresenta um comportamento oscilatório com um pico na região de OLR em torno de 2 g VS/(l.d) e um pico global na região de 6 g VS/(l.d). É importante notar que o substrato CD foi o que apresentou maior concentração volumétrica de metano, chegando a 85%, enquanto para FW e FVW o máximo foi atingido em torno de 70%, respectivamente. Portanto, essas informações são relevantes pois sugerem que existe uma faixa ideal de OLR para maximizar a produção de metano para cada tipo de substrato e que é importante considerar essa variável ao operar um reator anaeróbico, a fim de otimizar a produção de metano, corroborando o que foi discutido por Hagen et al. (2014), que discutiu a importância de determinar a taxa de carga ótima para cada tipo de resíduo e condições de operação. Era esperada uma maior influência do OLR na concentração volumétrica de metano, fato que não foi observado nas atuais superfícies. Devido a isso, uma análise mais detalhada dos dados para a confecção de superfícies de resposta mais específicas se faz necessário. Esse estudo será apresentado na seção a seguir. As Figuras 37, 38 e 39 apresentadas abaixo

apresentam as superfícies de resposta obtidas a partir da combinação entre temperatura e volume de reator.

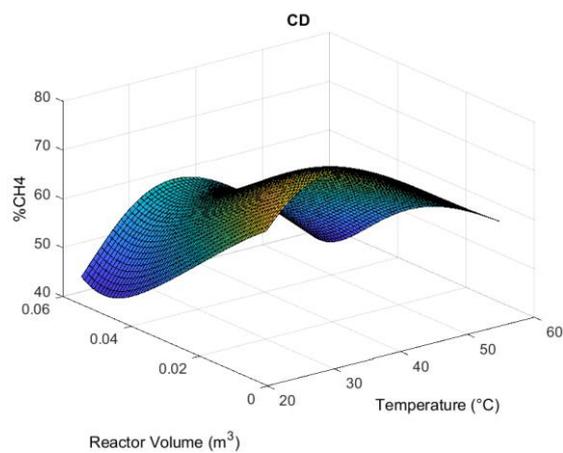


Figura 37 - Superfície de resposta para temperatura x volume de reator para CD, %CH4.

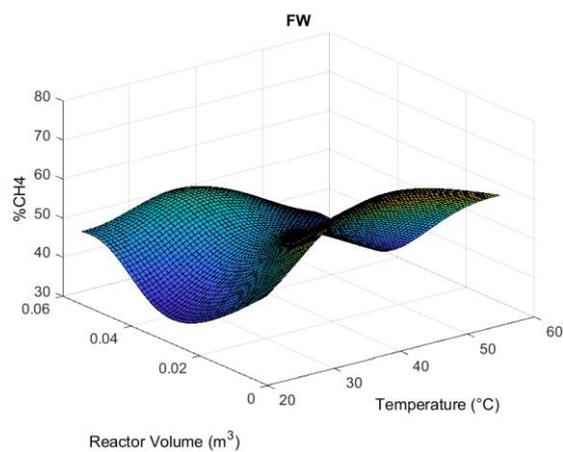


Figura 38 - Superfície de resposta para temperatura x volume de reator para FW, %CH4.

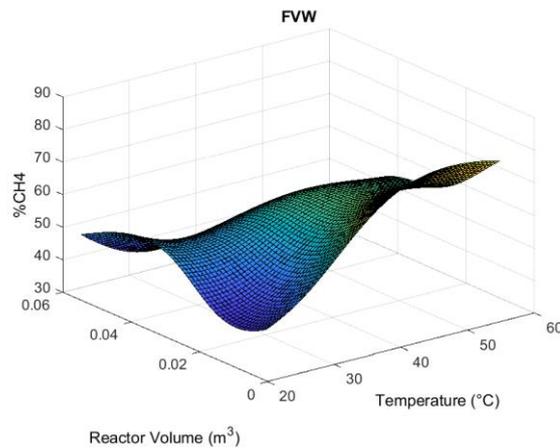


Figura 39 - Superfície de resposta para temperatura x volume de reator para FVW, %CH₄.

Observou-se que para os tipos de substratos FW e FVW avaliados na Figuras acima, a concentração volumétrica de metano foi maior na região termofílica de temperaturas, entre 40°C e 50°C. Esses resultados estão em consonância com a literatura, que aponta para uma produção maior de metano em temperaturas termofílicas. Estudo realizado por Min Ho Kim et al. (2017) mostrou que a utilização de condições termofílicas de temperatura (55°C) resultou em maiores taxas de produção de metano e rendimentos em comparação com condições mesofílicas (35°C). Além disso, Carlos Escamilla-Alvarado et al. (2010) identificaram que a produtividade de metano foi superior em regime termofílico. É importante observar também que para o substrato CD, a região de melhor desempenho foi atingida em temperaturas mesofílicas, entre 30°C e 40°C. Estudos (Smith e Almquist, 2014; Ros et al., 2013) indicam um aumento na concentração volumétrica de metano para a codigestão de substratos operando em temperaturas em torno de 35°C. Esses achados evidenciam a importância da temperatura na produção de metano em reatores anaeróbicos e reforçam a necessidade de avaliar a temperatura ideal para cada tipo de substrato e condição de operação do reator, a fim de maximizar a produção de metano.

A Figura 40 apresenta o gráfico que mostra a relação entre o tipo de substrato e o pH.

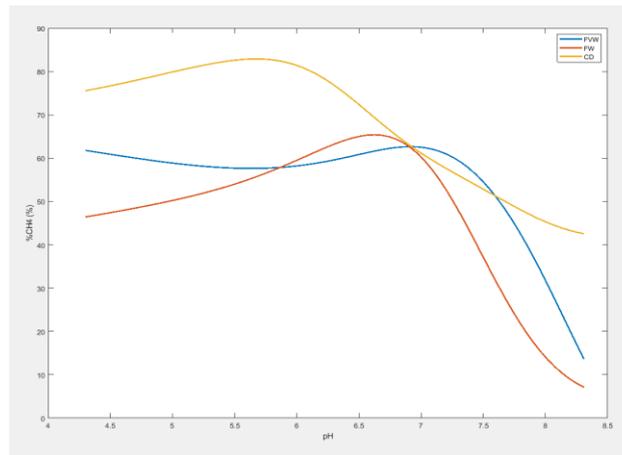


Figura 40 - Gráfico para tipo de substrato x pH, %CH₄.

É possível observar da Figura 45 que para os três tipos de substrato, há uma maior concentração volumétrica de metano nas condições mesofílicas de pH. Por outro lado, ocorre um decaimento na concentração volumétrica de metano para condições termofílicas de pH. Esse comportamento é mais acentuado para FW e FVW e mais suave para CD. Para CD, a região de maior concentração volumétrica de metano ocorre em valores de pH 5,4 e 6,2, enquanto que para FVW, ocorre entre pH 6,5 e 7,5 e para FW, pH entre 6 e 7. É importante destacar que há um decaimento na concentração volumétrica de metano à medida que o pH aumenta, o que é esperado, visto que na fase mesofílica de pH (baixo pH), ocorre a produção de moléculas de metano (ABEBE, 2017). Esses resultados estão em acordo com os demais resultados encontrados no presente estudo, onde a variável pH está envolvida. Portanto, a manutenção de um pH adequado é crucial para maximizar a produção de metano.

4.3.4 Avaliação dos modelos de produção de biogás e metano com o conjunto de dados completo

O modelo atual, apesar de fornecer valiosos *insights*, apresenta limitações notáveis, particularmente em condições extremas de pH e volumes elevados de reator, onde as variações se tornam imprecisas e extrapolaram. Além disso, a influência dos sólidos voláteis na produção de biogás permanece inconclusiva, dificultando uma compreensão abrangente do processo. Para superar essas

limitações, é imperativo direcionar esforços para estudos mais detalhados e específicos, adaptados a cada tipo de biomassa. Um modelo mais específico, estruturado considerando as características individuais de substratos como FW, FVW e CD, poderá oferecer resultados mais precisos e aplicáveis. Ao focar em modelos específicos por biomassa será possível desenvolver modelos mais refinados e contextualmente relevantes. Essa abordagem específica por biomassa é crucial para avançar nossa compreensão e otimizar a produção de biogás, proporcionando não apenas *insights* mais precisos, mas também orientando práticas operacionais mais eficientes e sustentáveis no campo da digestão anaeróbica. Os resultados desses novos modelos serão mostrados na seção a seguir.

4.4 Resultados dos modelos específicos por biomassa e comparações

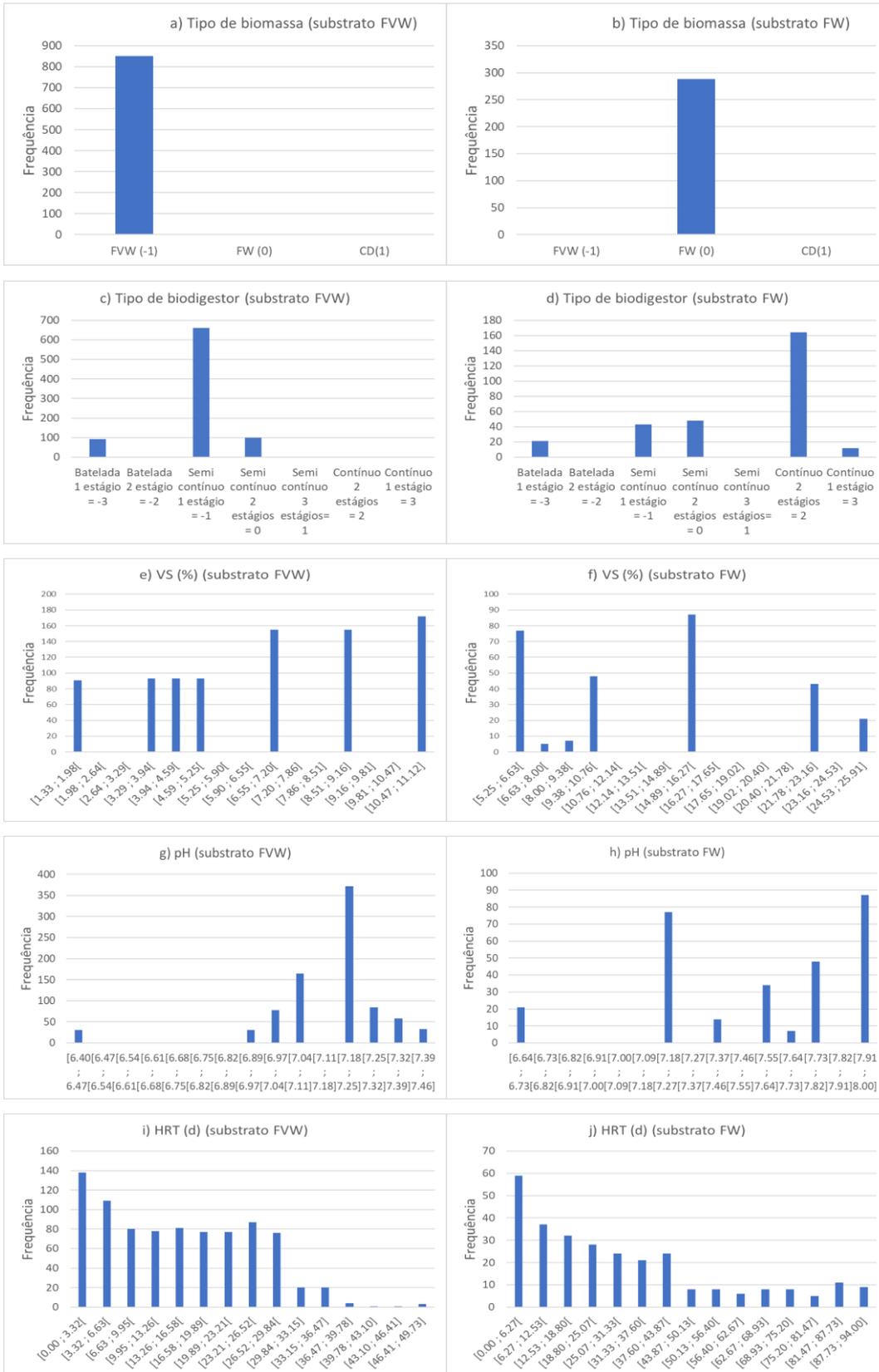
4.4.1 PAB - Banco de dados, modelos e superfícies

A partir do bando de dados PAB refinado, separou-se os dados em subconjuntos divididos pela biomassa. Cada biomassa (FW, FVW e CD) formou um novo conjunto de banco de dados denominado banco de dados PAB substrato FVW, PAB substrato FW e PAB substrato CD. A quantidade de elementos em cada banco de dados está descrita na Tabela 18. Devido a pouca quantidade de dados disponíveis para o subgrupo CD, decidiu-se por deixar esse subconjunto de fora das análises a seguir, uma vez que a pouca variedade de dados poderia afetar negativamente os resultados encontrados.

Tabela 19 - Quantidade de amostras em cada subconjunto de dados PAB. Fonte: Elaboração própria

Banco de dados	Amostras disponíveis
PAB substrato FVW	852 amostras
PAB substrato FW	288 amostras
PAB substrato CD	84 amostras

Os histogramas para os dois bancos de dados PAB (substrato FVW e substrato FW) são mostrados na Figura 41 a seguir.



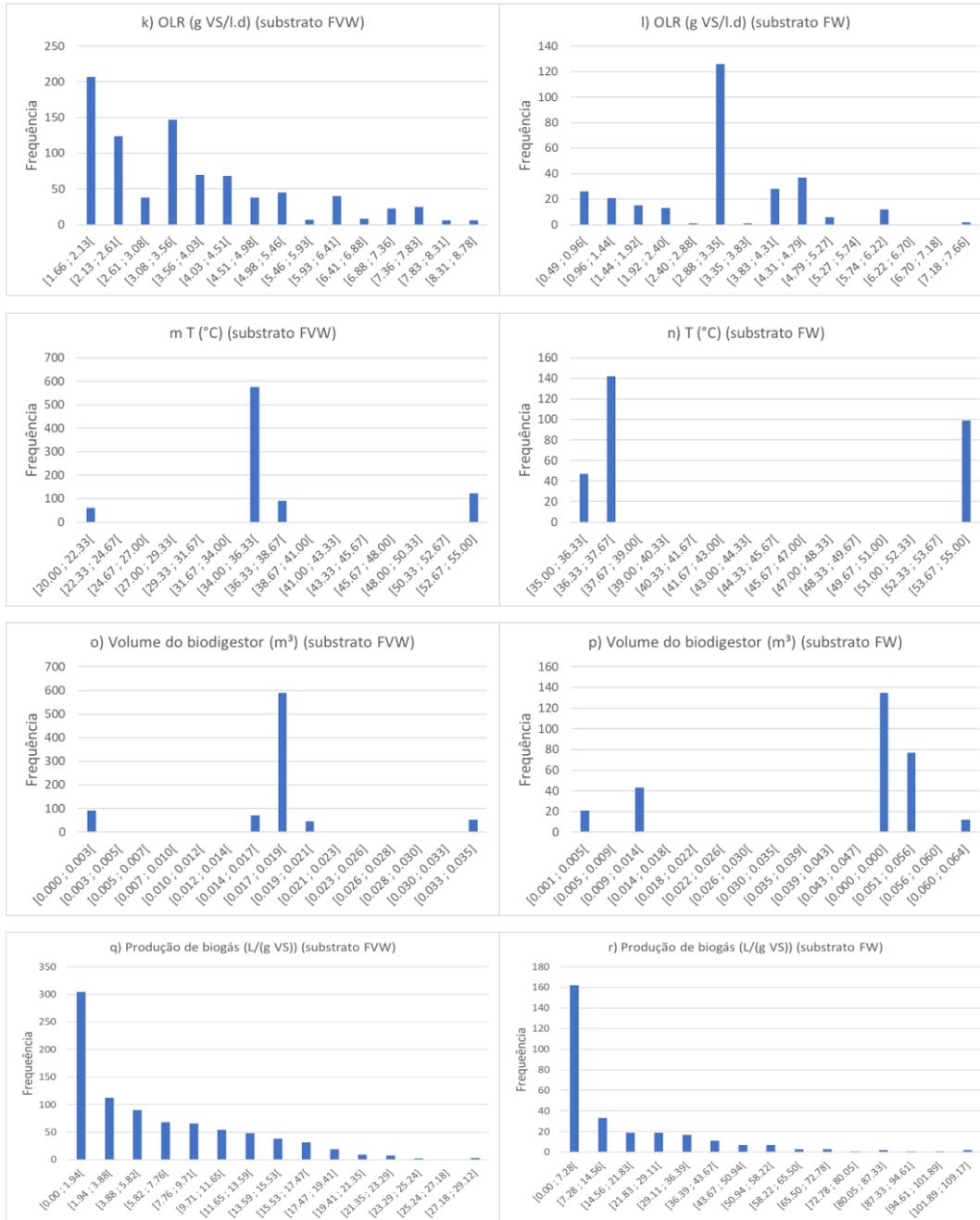


Figura 41 - Histogramas para os bancos de dados PAB por substrato: figuras das distribuições de “PAB substrato FVW” e “PAB substrato FW” para tipo de biomassa, tipo de reator, teor de sólidos voláteis (VS), pH, taxa de carga orgânica (OLR), tempo de retenção hidráulica (HRT), temperatura, volume do reator e produção acumulada de biogás.

Os histogramas gerados a partir dos dados PAB substrato FVW mostraram que existem regiões do banco de dados com nenhuma ocorrência de valores para algumas variáveis. Em particular, as regiões com tipo de reator batelada de 2 estágios, semi contínuo de 3 estágios, contínuo de 2 estágios e contínuo de 1 estágio.

Com relação ao teor de VS, os dados coletados estão apresentando uma distribuição quase homogênea na faixa de 1,33% a 11,12% e é importante notar que existem classes intermediárias vazias. O histograma para pH segue uma distribuição quase normal centrada na faixa de 7,18-7,25. Os valores de HRT concentram-se em uma duração de até 1 mês, conforme relatado na literatura. A faixa de OLR com maior população é 1,66-2,13 gVS/(l.d), mas valores tão altos quanto 8,78 gVS/(l.d) podem ser encontrados, porém raros. A distribuição da temperatura é majoritariamente unimodal com a maioria dos dados na temperatura mesófila (cerca de 35°C) e outras próximas dos 22°C (temperatura psicofílica) e dos 55°C (temperatura termofílica). O mesmo comportamento é observado para o histograma de volume de reator, onde reatores com volume de aproximadamente 18L são os mais representados, confirmando que os dados são coletados em escala laboratorial ou piloto. No histograma para produção acumulada de biogás os valores mais baixos ainda são os mais abundantes e as classes com valores mais altos ainda foram representadas.

Para o conjunto de dados PAB substrato FW o comportamento dos histogramas segue em grande parte os de PAB substrato FVW, porém, é importante destacar certas diferenças. Para PAB substrato FW o tipo de biodigestor que não está representado é batelada de 2 estágios e semi contínuo de 3 estágios. O histograma de VS tem classes bem representadas acima de 11%, a classe mais bem representada para volume de reator é em torno de 49 L enquanto que para o substrato FVW é 18L. Além disso para FW a produção acumulada de biogás atinge maiores valores, atingindo pico de 109,17 L/gVS enquanto para FVW o pico é de 29,12 L/gVS.

Seguindo o fluxo de trabalho definido na metodologia, a segunda etapa, etapa de desenvolvimento das redes, os resultados são mostrados a seguir para PAB substrato FVW e PAB substrato FW. A Tabela com configuração das melhores redes é apresentada na Tabela 19 abaixo. Desse modo, é possível observar da Tabela 19 que as 3 melhores topologias de RNAs identificadas tiveram desempenho similar, tendo a topologia do banco PAB substrato FVW (*trainbr_logsig_11_purelin_1*) apresentado um valor de SSE maior que as demais, de 2,2187. Porém ainda é considerado um valor baixo, indicando baixo grau de variabilidade dentro do conjunto de dados.

Tabela 20 - Comparação entre as configurações das melhores redes PAB. Fonte: Elaboração própria.

Banco de dados	Melhor topologia	SSE	R2 treino	R2 teste
PAB substrato FVW	<i>trainbr_logsig_11_purelin_1</i>	2,2187	0,9741	0,9613
PAB substrato FW	<i>trainlm_tansig_9_purelin_1</i>	0,0029	0,9999	0,9996
PAB refinado completo	<i>trainbr_tansig_11_purelin_1</i>	0,1808	0,9950	0,9938

A Figura 42 abaixo apresenta os resultados de desempenho para as novas topologias de redes neurais.

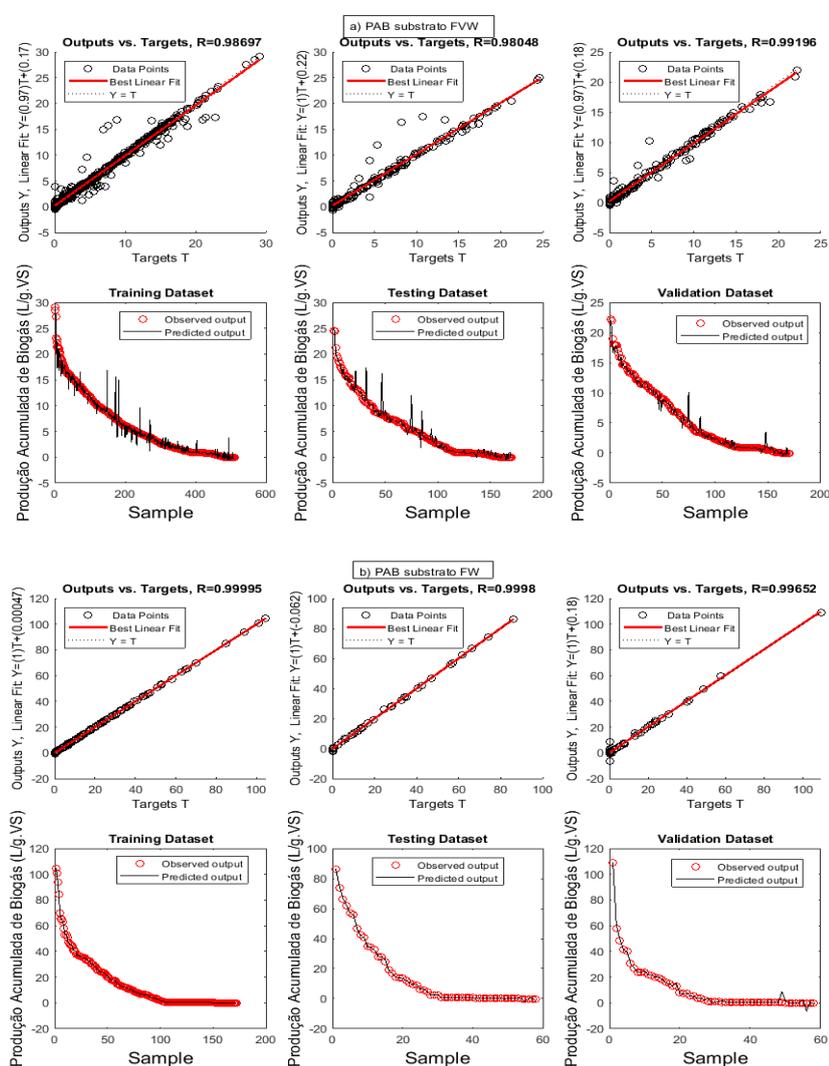


Figura 42 – a) Resultados para treinamento, teste e validação da RNA PAB substrato FVW, b) Resultados para treinamento, teste e validação da RNA PAB substrato FW. Fonte: Elaboração própria.

Como observado na Figura 42, é possível considerar que os modelos novos são capazes de prever bem o volume acumulado de biogás produzido para os insumos considerados, obtendo um R^2 igual a 0,9722 e 0,9980 para PAB substrato FVW e PAB substrato FW, respectivamente, quando aplicado a RNA no banco de dados inteiro. Esses valores são bem próximos do R^2 encontrado pelo modelo anterior utilizando o banco de dados completo PAB refinado, que foi de 0,9938.

Na terceira etapa do fluxo de trabalho, é apresentada as superfícies de resposta. A Figura 43 apresenta as comparações entre as superfícies encontradas na etapa anterior e as novas superfícies de resposta utilizando as topologias novas usando os bancos de dados PAB substrato FVW e PAB substrato FW. Os valores fixos das variáveis e os valores de máximos e mínimos foram mantidos iguais aos do banco de dados PAB refinado a fim de possibilitar a comparação justa entre os resultados. Desse modo, os valores das Tabelas 12 e 13 se mantêm.

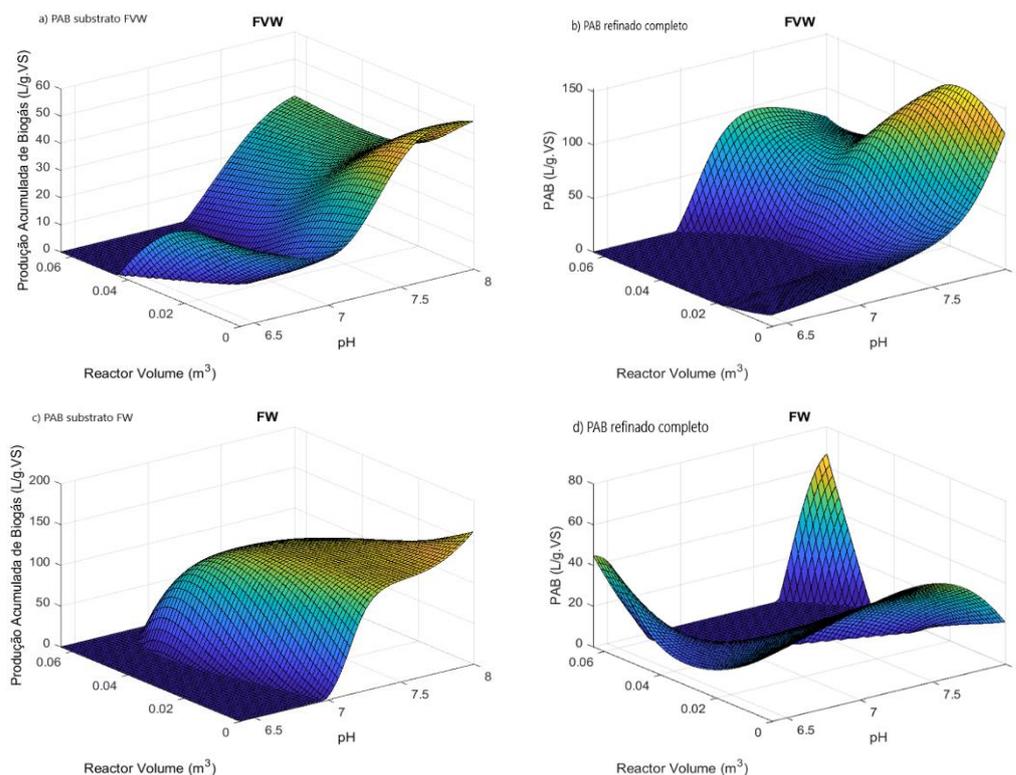


Figura 43 – Comparação entre as superfícies de resposta para a combinação entre pH x volume de reator para PAB. A) Superfície usando banco PAB substrato FVW, B) Superfície usando banco PAB refinado completo, C) Superfície usando banco PAB substrato FW, D) Superfície usando banco PAB refinado completo.

Da Figura 43 é possível observar que o uso dos bancos de dados específicos por biomassa (banco PAB substrato FVW e PAB substrato FW) resultou em superfícies mais regulares, principalmente para o substrato FW. Para o substrato FVW (a) a máxima produção acumulada de biogás continua sendo atingida para valores de pH termofílicos, em torno de 8, comportamento esse que foi seguido para a nova superfície de resposta para FW (c). É importante notar que a máxima produção de biogás para a superfície (a) é menor do que a atingida para a superfície (b). As novas RNAs treinadas apresentaram resultados consistentes para FVW e FW, seguindo as análises obtidas na sessão 4.3.1 do trabalho.

A seguir é apresentada as figuras para temperatura versus volume de reator, seguindo a mesma sistemática adotada anteriormente.

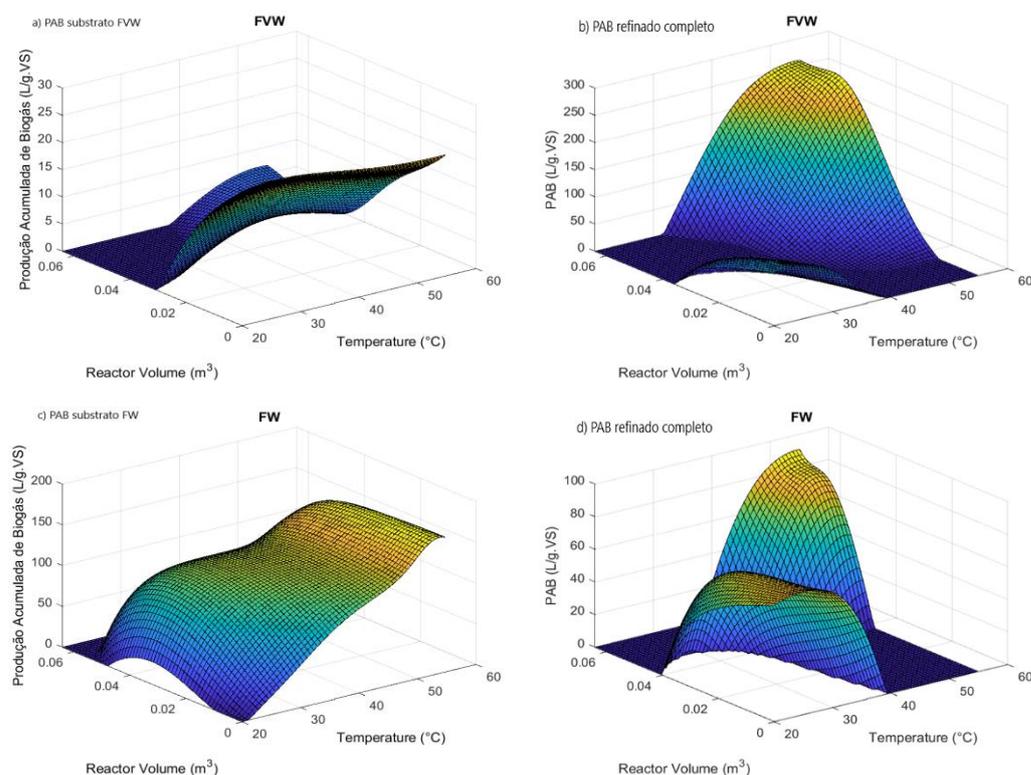


Figura 44 - Comparação entre as superfícies de resposta para a combinação entre temperatura x volume de reator para PAB. A) Superfície usando banco PAB substrato FVW, B) Superfície usando banco PAB refinado completo, C) Superfície usando banco PAB substrato FW, D) Superfície usando banco PAB refinado completo.

A Figura 44 mostra que as superfícies ficaram mais regulares com os bancos de dados específicos por biomassa (banco PAB substrato FVW e PAB substrato

FW). Para o substrato FVW (a) a máxima produção acumulada de biogás está agora dentro dos limites observados no banco de dados através dos histogramas da Figura 45 (q). A nova superfície para o substrato FW apresentou um valor máximo de PAB maior do que o presente no banco de dados (109,17 L/g.VS) para temperaturas acima de 45°C, mostrando assim uma possível limitação do modelo para valores de temperatura acima desse valor. De modo geral, as novas superfícies mantêm valores coerentes para PAB na faixa de temperatura mesofílica, indicando assim um ponto forte do modelo para essa faixa de temperatura (entre 20°C e 45°C). Desse modo, os resultados para FVW e FW foram consistentes com as novas RNAs treinadas.

As figuras para a combinação de variáveis VS versus volume de reator são apresentadas a seguir:

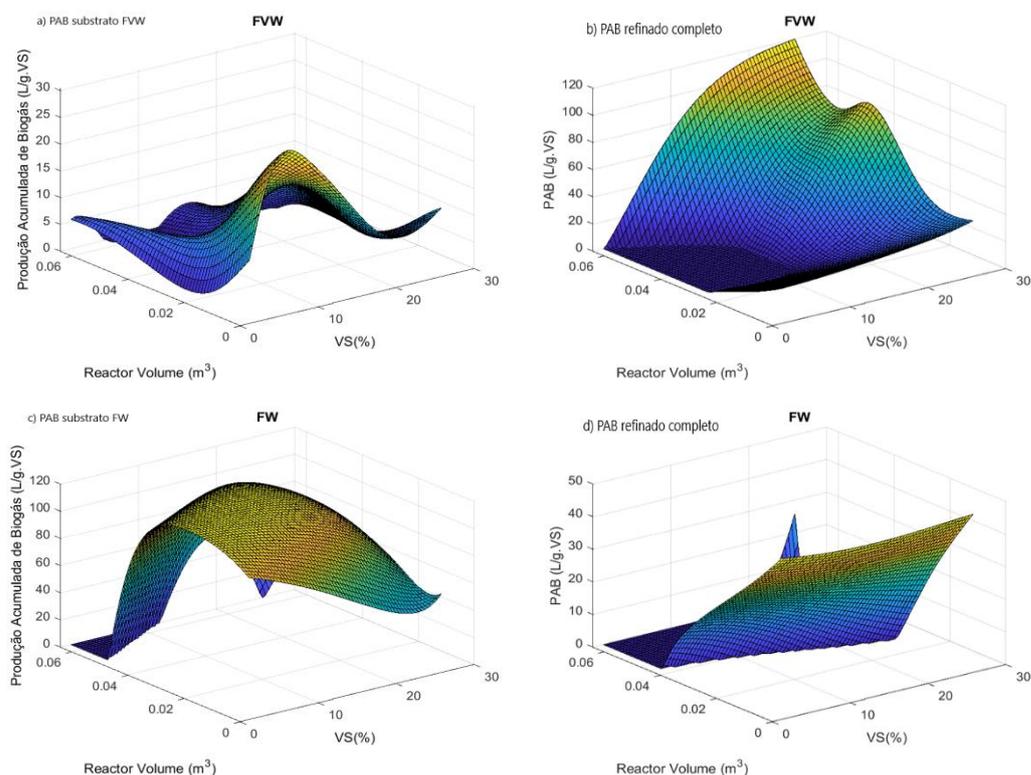


Figura 45 - Comparação entre as superfícies de resposta para a combinação entre VS x Volume de reator para PAB. A) Superfície usando banco PAB substrato FVW, B) Superfície usando banco PAB refinado completo, C) Superfície usando banco PAB substrato FW, D) Superfície usando banco PAB refinado completo.

A Figura 45 mostra que os novos bancos de dados específicos por biomassa permitiram uma maior variação de PAB em função do VS, que não era tão evidente

nas superfícies anteriores. Para ambos os substratos FVW e FW, o VS ótimo foi de cerca de 10%, conforme as superfícies geradas. Na figura (a), o PAB máximo foi obtido para volumes de reator menores que 20L, enquanto na figura (c), os melhores resultados foram na faixa de volume de reator entre 20L e 40L. Os PAB máximos para as duas novas superfícies estavam dentro dos limites dos bancos de dados correspondentes, indicando uma boa convergência dos resultados da rede neural.

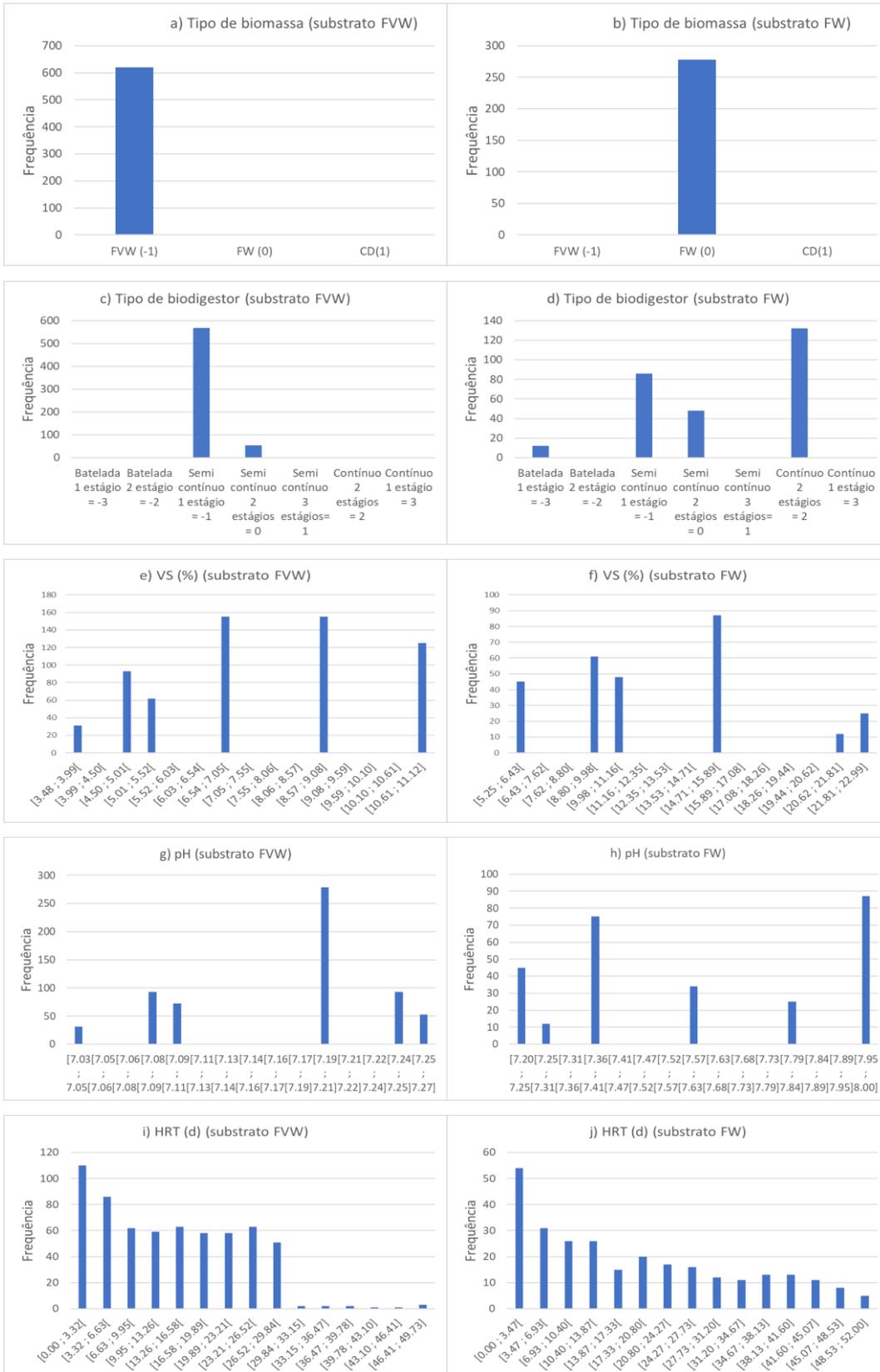
4.4.2 PAM - Banco de dados, modelos e superfícies

Seguindo o que foi feito na seção 4.4.1, a partir do banco de dados PAM refinado, os dados foram separados em subconjuntos com base na biomassa (FW, FVW e CD), criando três novos conjuntos de dados: PAM substrato FVW, PAM substrato FW e PAM substrato CD. O número de elementos em cada conjunto está detalhado na Tabela 20. Assim como ocorreu em PAB, devido à escassez de dados no subgrupo CD, optou-se por excluí-lo das análises subsequentes, pois a falta de variedade nos dados poderia impactar negativamente os resultados.

Tabela 21 - Quantidade de amostras em cada subconjunto de dados PAM. Fonte: Elaboração própria

Banco de dados	Amostras disponíveis
PAB substrato FVW	<i>621 amostras</i>
PAB substrato FW	<i>278 amostras</i>
PAB substrato CD	<i>91 amostras</i>

Os histogramas dos dois conjuntos de dados PAM (um com substrato FVW e outro com substrato FW) estão apresentados na Figura 46 abaixo.



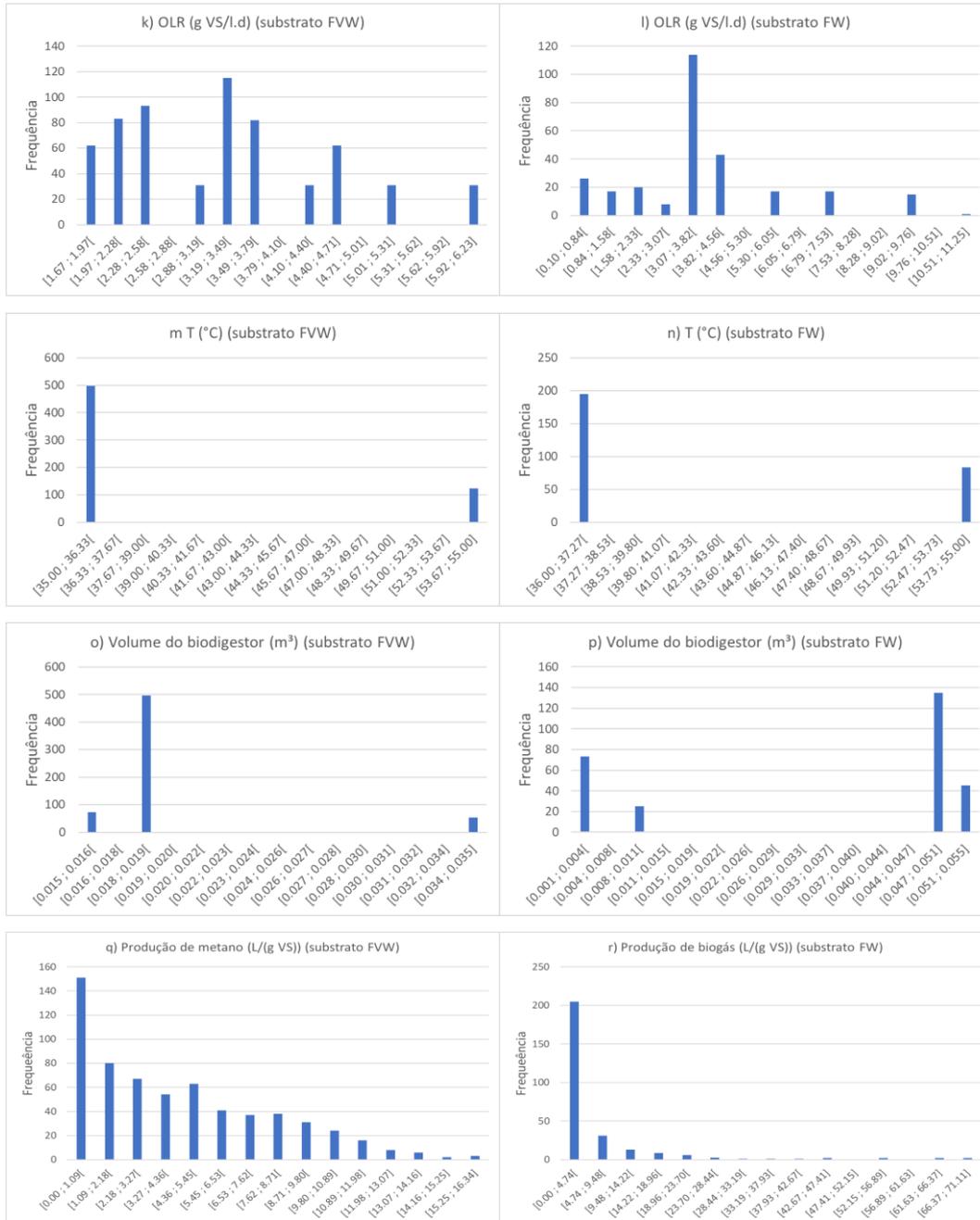


Figura 46 - Histogramas para os bancos de dados PAM por substrato: figuras das distribuições de “PAM substrato FVW” e “PAM substrato FW” para tipo de biomassa, tipo de reator, teor de sólidos voláteis (VS), pH, taxa de carga orgânica (OLR), tempo de retenção hidráulica (HRT), temperatura, volume do reator e produção acumulada de biogás.

Os histogramas derivados dos dados do substrato FVW revelaram áreas em branco, indicando ausência de dados em algumas variáveis. Especificamente, estas áreas correspondem aos tipos de reatores, tais como batelada de 1 estágio, batelada de 2 estágios, semi contínuo de 3 estágios, contínuo de 2 estágios e contínuo de 1

estágio, bem como temperaturas variando de 36,33°C a 53,67°C e volumes do reator situados entre 19L e 35L. É notável que os valores de pH estão restritos a uma faixa estreita, variando de 7,03 a 7,27. É importante salientar que, apesar da homogeneidade na distribuição dos dados nas classes, existem intervalos intermediários vazios nos histogramas de VS e OLR. Ademais, os valores de HRT raramente ultrapassam um mês, conforme corroborado pela literatura. No histograma da produção acumulada de metano, é observado que os valores mais baixos predominam, enquanto classes com valores mais elevados também estão representadas, indicando uma tendência específica na distribuição desses dados.

No contexto dos dados do substrato FW no conjunto PAM, os histogramas exibem padrões semelhantes aos observados no substrato FVW do mesmo conjunto. Contudo, é fundamental notar algumas discrepâncias. Para o substrato FW no conjunto PAM, não há representação nos tipos de biodigestores em batelada de 2 estágios, semi contínuo de 3 estágios e contínuo de 1 estágio. O histograma de VS revela classes bem definidas acima de 11%, com a classe mais frequente para o volume do reator situada em torno de 49 L, enquanto, para o substrato FVW, esse valor é de 18 L. Além disso, no caso do substrato FW, a produção acumulada de metano atinge valores mais elevados, chegando a um pico de 71,11 L/gVS, ao passo que para o substrato FVW, o pico é de 16,34 L/gVS. Essas variações sugerem diferenças significativas nos padrões de comportamento dos dados entre os substratos FW e FVW no contexto do conjunto de dados PAM.

A tabela que detalha a configuração das redes mais eficazes está disponível na Tabela 21 abaixo.

Tabela 22 - Comparação entre as configurações das melhores redes PAM. Fonte: Elaboração própria.

Banco de dados	Melhor topologia	SSE	R2 treino	R2 teste
PAM substrato FVW	<i>trainlm_tansig_11_purelin_1</i>	1,9488	0,9750	0,9844
PAM substrato FW	<i>trainlm_tansig_8_purelin_1</i>	0,0044	0,9997	0,9994
PAM refinado completo	<i>trainbr_logsig_8_purelin_1</i>	0,1089	0,9958	0,9859

Desse modo, é possível observar da Tabela 21 que as topologias para PAM substrato FW e PAM refinado completo tiveram desempenho similar, tendo a

topologia do banco PAM substrato FVW (*trainlm_tansig_11_purelin_1*) apresentado um desempenho pior tanto de R^2 quanto de SSE. Porém o resultado para essa RNA ainda é considerado um desempenho bom com R^2 próximo de 1 e SSE baixo, indicando baixo grau de variabilidade dentro do conjunto de dados. A Figura 47 abaixo apresenta os resultados de desempenho para as novas topologias de redes neurais.

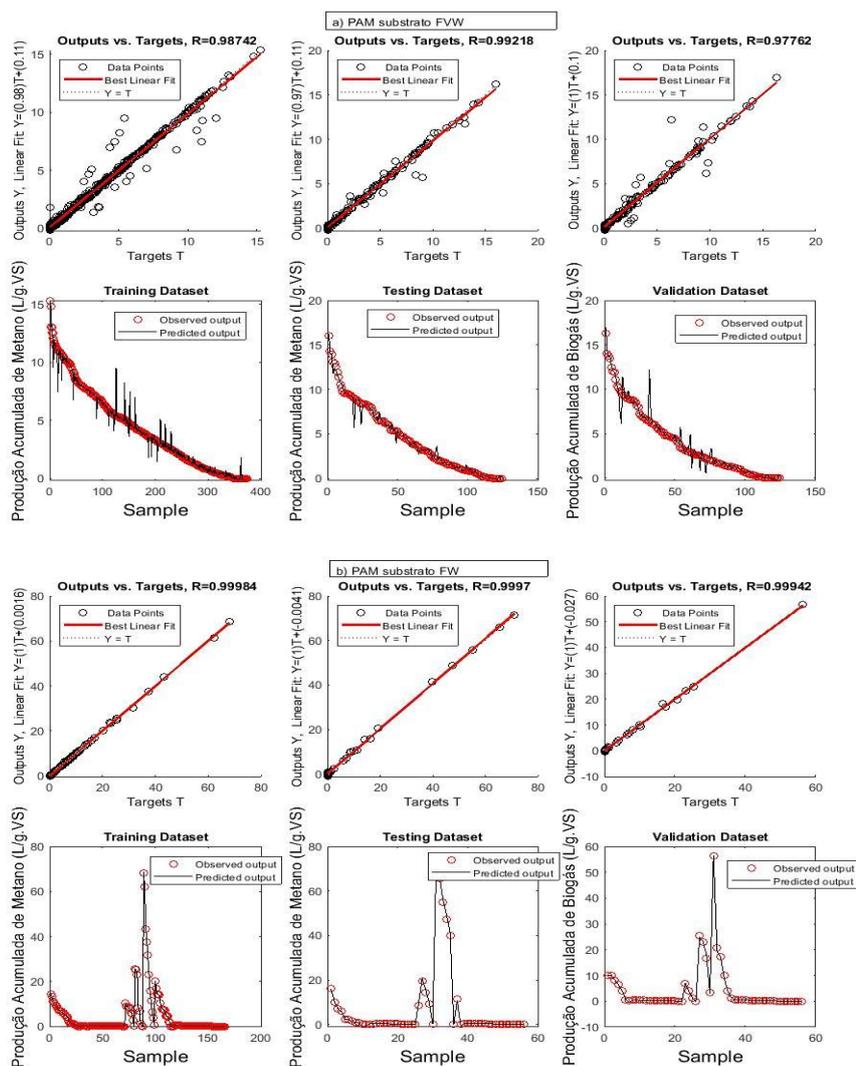


Figura 47 - a) Resultados para treinamento, teste e validação da RNA PAM substrato FVW, b) Resultados para treinamento, teste e validação da RNA PAM substrato FW. Fonte: Elaboração própria.

Conforme ilustrado na Figura 47, pode-se concluir que os novos modelos têm uma capacidade significativa de prever o volume acumulado de metano produzido para os insumos analisados. Eles alcançaram coeficientes de determinação (R^2) de 0,9724 e 0,9994 para PAM substrato FVW e PAM substrato

FW, respectivamente, quando aplicados a toda a extensão do banco de dados. Estes valores são próximos ao R^2 obtido pelo modelo anterior utilizando o banco de dados completo PAM refinado, que foi de 0,9926.

A Figura 48 compara as superfícies encontradas na fase anterior com as novas superfícies de resposta utilizando as topologias recentemente desenvolvidas, empregando os conjuntos de dados PAM substrato FVW e PAM substrato FW. Para assegurar uma comparação justa entre os resultados, os valores das variáveis fixas, bem como os valores máximos e mínimos, permaneceram os mesmos do banco de dados PAM refinado.

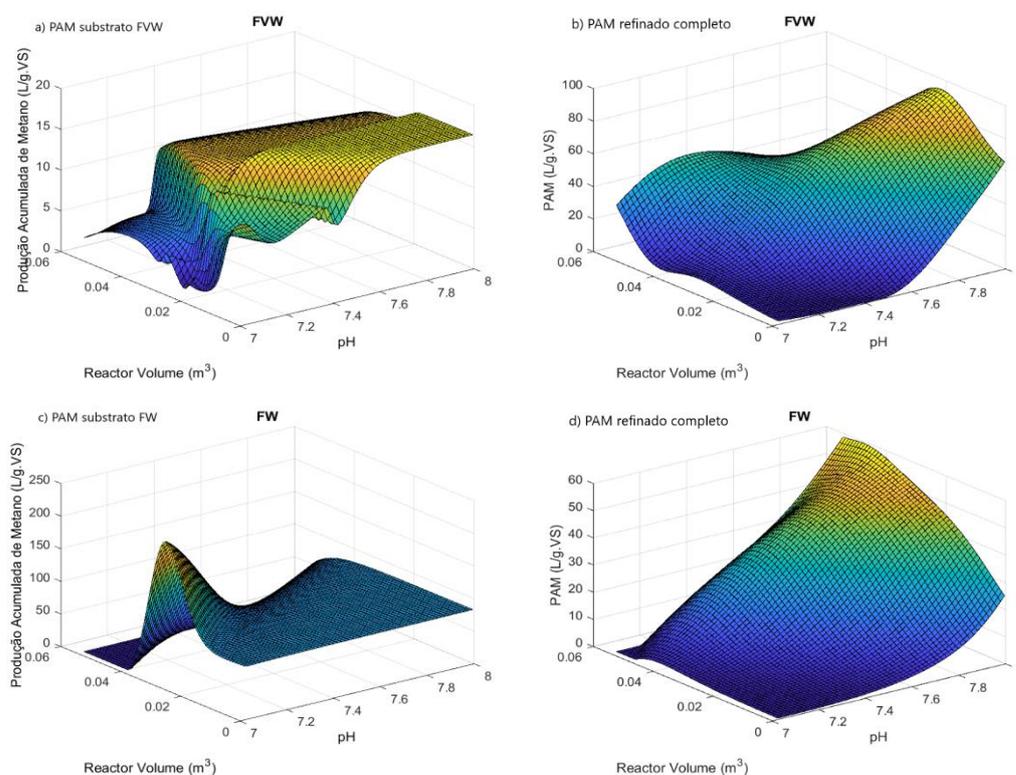


Figura 48 - Comparação entre as superfícies de resposta para a combinação entre pH x volume de reator para PAM. A) Superfície usando banco PAM substrato FVW, B) Superfície usando banco PAM refinado completo, C) Superfície usando banco PAM substrato FW, D) Superfície usando banco PAM refinado completo.

A análise da Figura 48 revela que a utilização de bancos de dados específicos por biomassa (ou seja, PAM substrato FVW e PAM substrato FW) não resultou em superfícies de resposta superiores às encontradas anteriormente. Estes modelos não conseguiram prever com consistência a influência das variáveis pH e volume do reator na produção acumulada de metano, indicando limitações

significativas. Esta limitação pode ser atribuída às particularidades dos novos conjuntos de dados, que apresentam valores de pH restritos a uma estreita faixa, como evidenciado nos histogramas da Figura 51. Estes resultados indicam uma possível restrição do modelo em lidar com valores fora da gama disponível nos bancos de dados.

A seguir é apresentada as figuras para VS versus volume de reator.

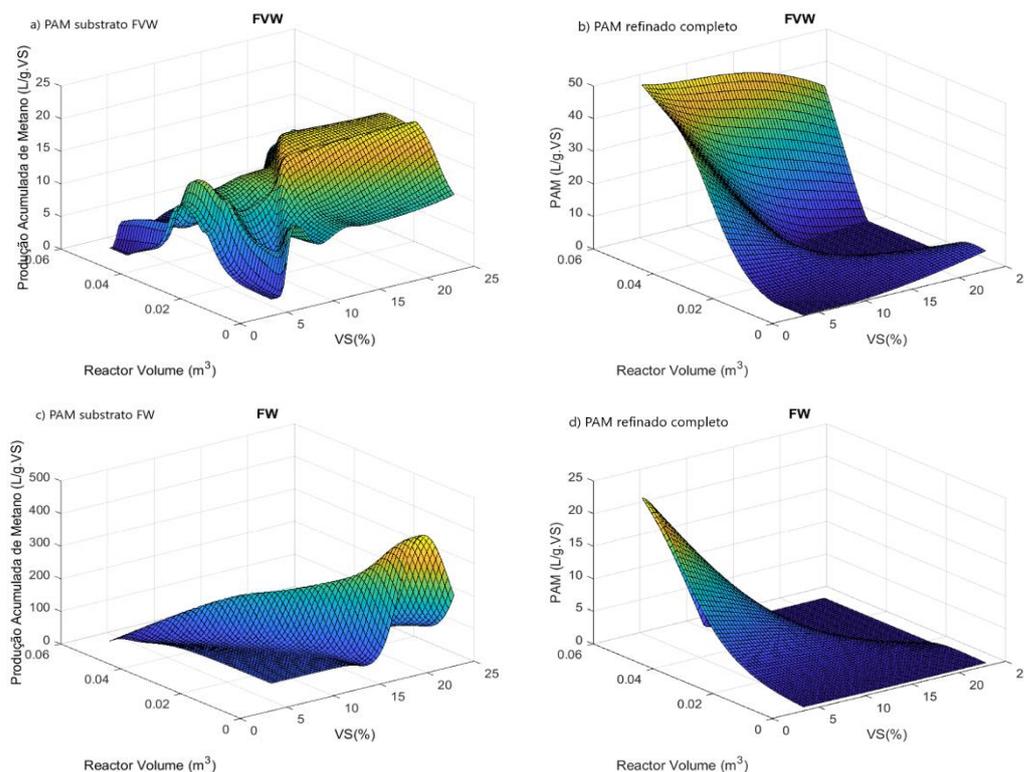


Figura 49 - Comparação entre as superfícies de resposta para a combinação entre VS x volume de reator para PAM. A) Superfície usando banco PAM substrato FVW, B) Superfície usando banco PAM refinado completo, C) Superfície usando banco PAM substrato FW, D) Superfície usando banco PAM refinado completo.

A Figura 49 evidencia uma notável irregularidade nas superfícies, especialmente para o cenário do PAM substrato FVW. No caso do PAM substrato FW, a utilização do banco de dados mais restrito resultou em valores extrapolados para a produção acumulada de metano, ultrapassando 100 L/g.VS. Nesse contexto, os modelos recém-desenvolvidos a partir dos bancos de dados específicos por biomassa, notadamente no cenário PAM, demonstraram uma deterioração na precisão das previsões por meio das superfícies de resposta. Essa deterioração pode

ser atribuída à redução do tamanho do banco de dados, resultando em dados mais restritos e limitados em certas faixas. A restrição das variáveis e a diminuição do espectro de dados disponíveis podem ter exacerbado a complexidade do modelo, levando a uma representação menos precisa das interações entre as variáveis envolvidas. É essencial considerar essas limitações ao interpretar os resultados e, possivelmente, explorar estratégias para mitigar o impacto da redução do banco de dados na qualidade das previsões.

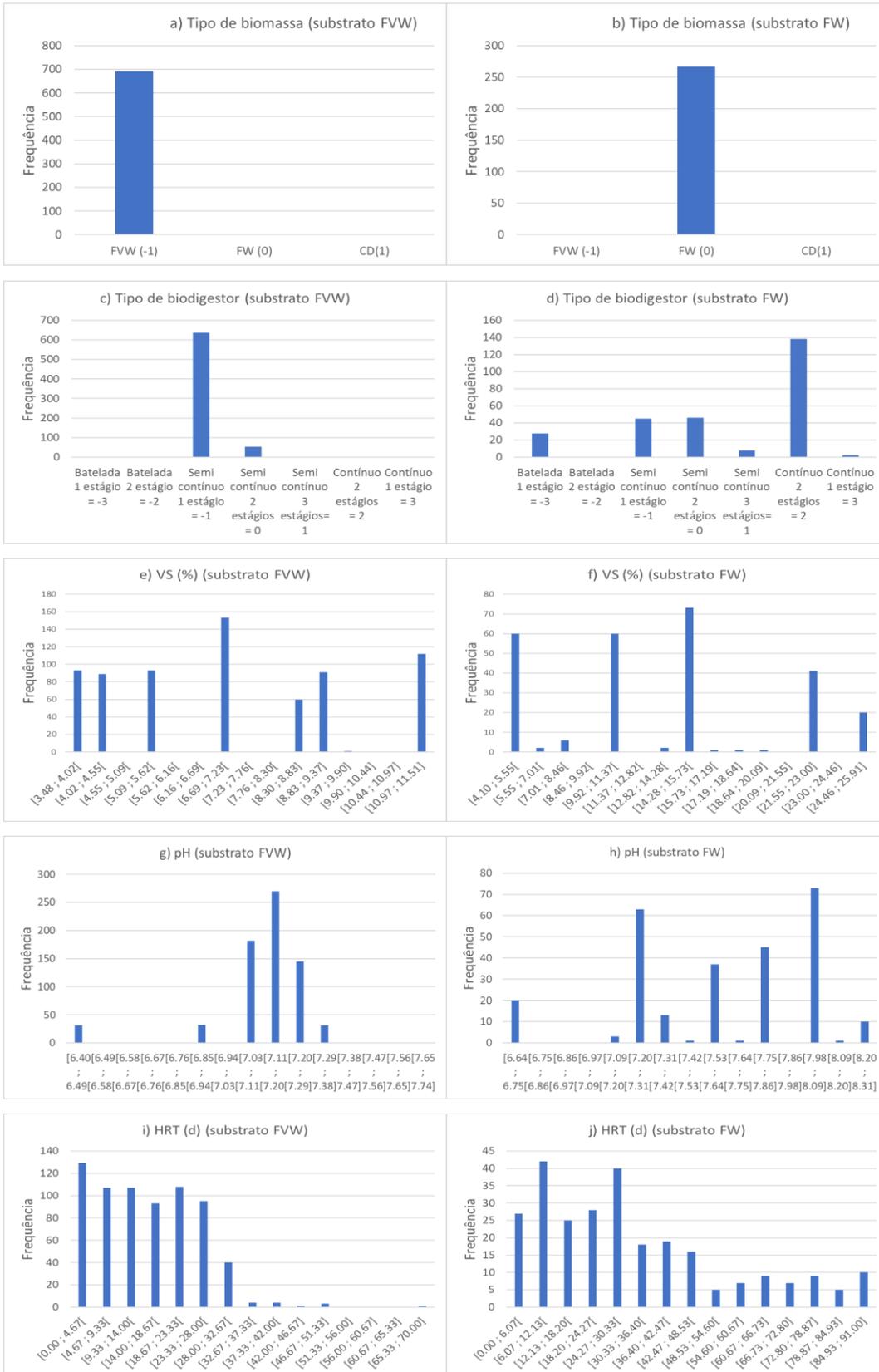
4.4.3 %CH₄ - Banco de dados, modelos e superfícies

Conforme descrito na seção 4.4.1, os dados foram novamente subdivididos com base na biomassa (FW, FVW e CD), originando três novos conjuntos de dados: %CH₄ substrato FVW, %CH₄ substrato FW e %CH₄ substrato CD. Os detalhes sobre o número de elementos em cada conjunto podem ser encontrados na Tabela 22. Da mesma forma que foi feito para as demais saídas, o subconjunto CD foi excluído das análises posteriores devido à falta de dados variados, que poderiam impactar negativamente nos resultados.

Tabela 23 - Quantidade de amostras em cada subconjunto de dados %CH₄. Fonte: Elaboração própria.

Banco de dados	Amostras disponíveis
%CH ₄ substrato FVW	<i>692 amostras</i>
%CH ₄ substrato FW	<i>267 amostras</i>
%CH ₄ substrato CD	<i>103 amostras</i>

Os histogramas dos dois conjuntos de dados PAM, um com substrato FVW e outro com substrato FW, são ilustrados na Figura 50 abaixo.



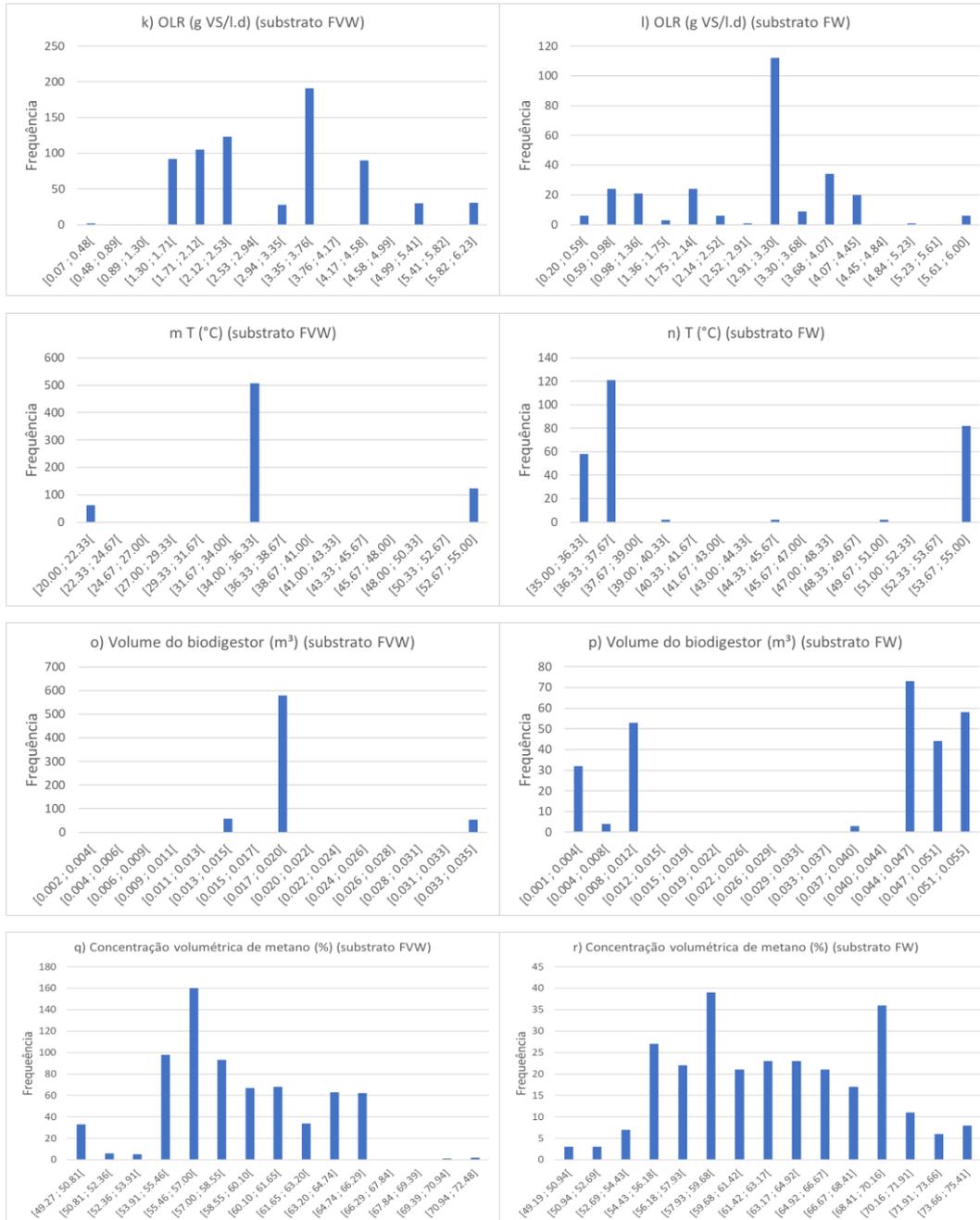


Figura 50 - Histogramas para os bancos de dados %CH₄ por substrato: figuras das distribuições de “%CH₄ substrato FWW” e “%CH₄ substrato FW” para tipo de biomassa, tipo de reator, teor de sólidos voláteis (VS), pH, taxa de carga orgânica (OLR), tempo de retenção hidráulica (HRT), temperatura, volume do reator e produção acumulada de biogás.

Os histogramas derivados dos dados do substrato FWW revelaram lacunas, indicando a falta de informações em diversas variáveis. Especificamente, essas lacunas correspondem aos tipos de reatores, como batelada de 1 estágio, batelada de 2 estágios, semi contínuo de 3 estágios, contínuo de 2 estágios e contínuo de 1

estágio, além de algumas categorias de VS, pH e OLR. Além disso, os valores de HRT raramente excedem um mês, conforme corroborado pela literatura. O histograma da temperatura mostra predominantemente uma distribuição unimodal, com valores concentrados na faixa de 34°C a 36,33°C; entretanto, é possível encontrar valores tão baixos quanto 20°C e tão altos quanto 55°C. Similarmente, o volume do reator segue a mesma tendência, com a faixa principal situada entre 17L e 20L.

Dentro do contexto dos dados do substrato FW no conjunto de %CH₄, os histogramas demonstram padrões semelhantes aos observados no substrato FVW do mesmo conjunto. No entanto, é crucial notar algumas discrepâncias. Para o substrato FW, a ausência de representação ocorre apenas nos tipos de biodigestores em batelada de 2 estágios. O histograma de VS exibe classes claramente definidas acima de 11%, e há uma quantidade significativa de dados para HRT acima de 30.

A configuração das redes mais eficazes está disponível na Tabela 23 abaixo.

Tabela 24 - Comparação entre as configurações das melhores redes %CH₄. Fonte: Elaboração própria.

Banco de dados	Melhor topologia	SSE	R² treino	R² teste
%CH ₄ substrato FVW	<i>trainbr_logsig_10_purelin_1</i>	3,6019	0,9235	0,8689
%CH ₄ substrato FW	<i>trainbr_tansig_10_purelin_1</i>	5,3501	0,8367	0,8448
%CH ₄ refinado completo	<i>trainbr_logsig_10_purelin_1</i>	11,450	0,9011	0,8460

Dessa forma, pode-se notar, a partir da Tabela 23, que as topologias para o %CH₄ no substrato FW e na %CH₄ do banco de dados refinado completo apresentaram desempenho semelhante, com a topologia do substrato FW (*trainbr_tansig_10_purelin_1*) exibindo um desempenho inferior em relação aos valores de R² de treino e teste. É relevante destacar que os valores de SSE para os conjuntos de dados restritos por biomassa diminuíram em comparação com o banco de dados completo. Esse padrão era esperado, dado o formato dos novos conjuntos de dados, que incluem informações que variam menos. A Figura 51 abaixo apresenta os resultados de desempenho para as novas topologias de redes neurais.

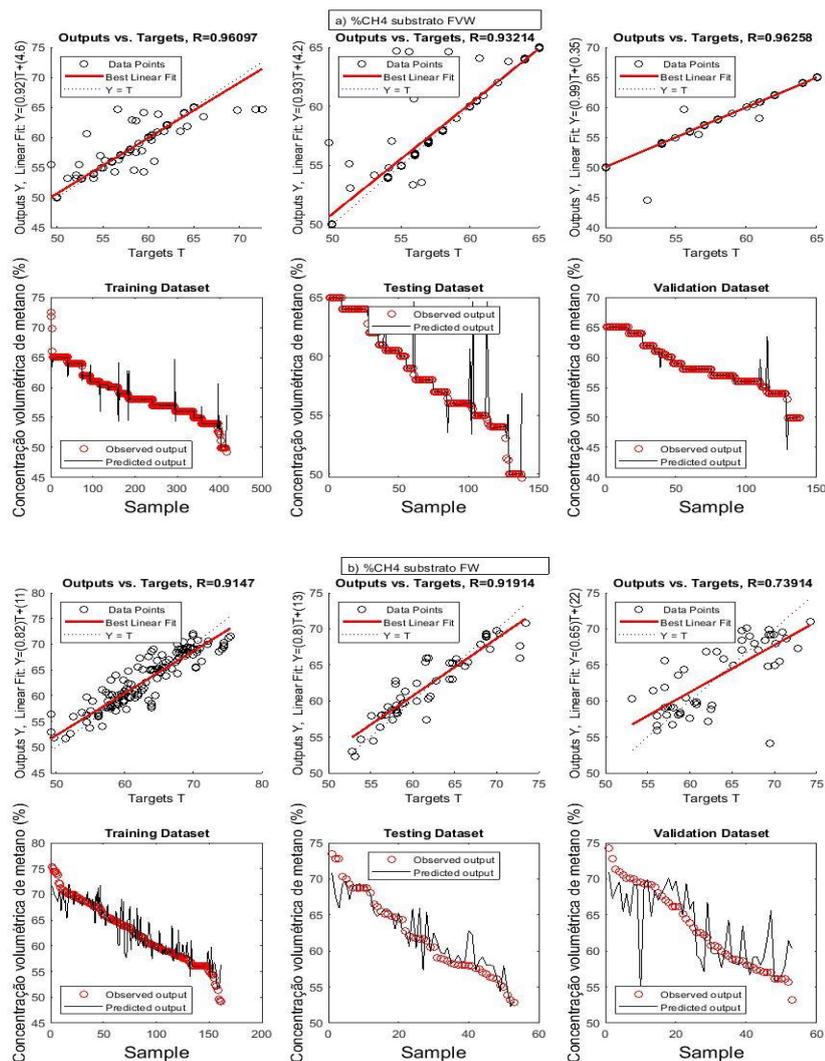


Figura 51 - a) Resultados para treinamento, teste e validação da RNA %CH₄ substrato FVW, b) Resultados para treinamento, teste e validação da RNA %CH₄ substrato FW. Fonte: Elaboração própria.

Conforme ilustrado na Figura 51, pode-se concluir que o novo modelo para o substrato FVW têm uma capacidade significativa de prever a concentração volumétrica de metano produzido para os insumos analisados, alcançando um coeficiente de determinação (R^2) de 0,9097, valor próximo ao R^2 obtido pelo modelo anterior utilizando o banco de dados completo %CH₄ refinado, que foi de 0,8809. Já para o substrato FW, o coeficiente de determinação foi de 0,7802 quando aplicado a toda a extensão do banco de dados novo, obtendo um desempenho inferior aos demais modelos. A análise do gráfico também revela que o desempenho na validação foi inferior devido a uma correlação menor ($R=0,7394$) em comparação com os conjuntos de treinamento e teste. Isso indica que o modelo pode

não estar generalizando adequadamente para novos dados, o que sugere a possibilidade de *overfitting*. A análise dos bancos de dados revela que o conjunto de dados para FW é de menor tamanho e menos diversificado em comparação com o banco de dados FVW. Portanto, um conjunto de dados de treinamento limitado pode não refletir adequadamente a diversidade da população total, resultando em um modelo que não generaliza eficientemente. Ademais, a presença de características redundantes pode induzir o modelo a aprender padrões que não são pertinentes para a previsão em novos dados, o que pode explicar o desempenho inferior do modelo para FW.

A Figura 52 compara as superfícies encontradas na fase anterior com as novas superfícies de resposta utilizando as topologias recentemente desenvolvidas, empregando os conjuntos de dados %CH₄ substrato FVW e %CH₄ substrato FW. Para assegurar uma comparação justa entre os resultados, os valores das variáveis fixas, bem como os valores máximos e mínimos, permaneceram os mesmos do banco de dados %CH₄ refinado.

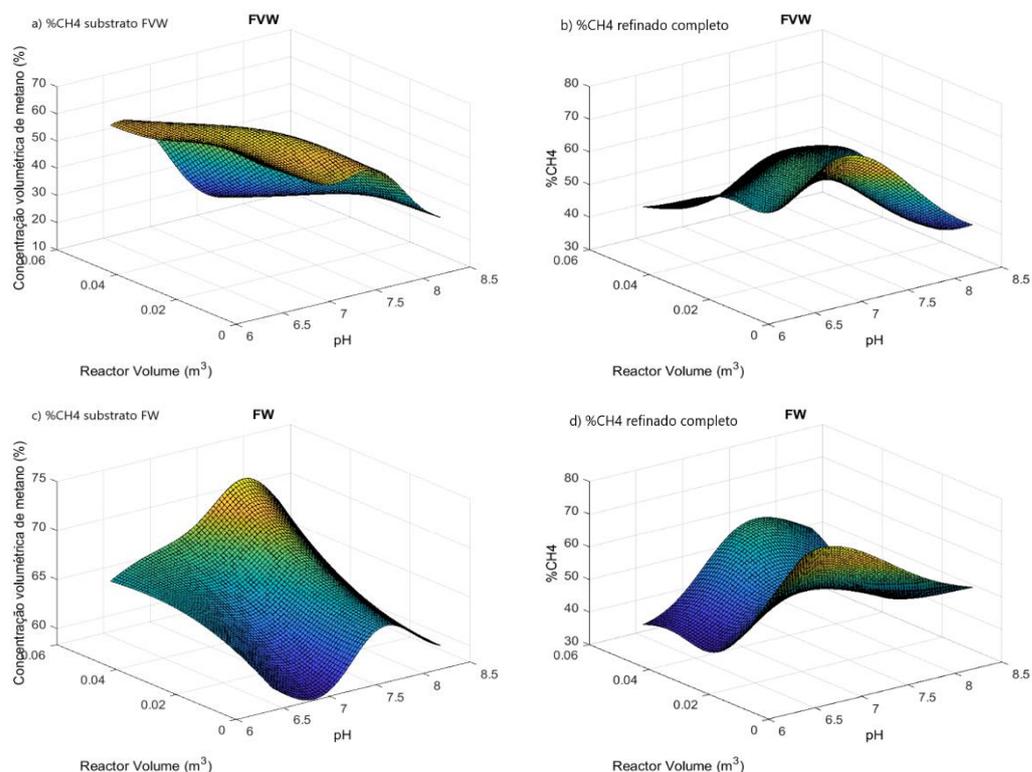


Figura 52 - Comparação entre as superfícies de resposta para a combinação entre pH x volume de reator para %CH₄. A) Superfície usando banco %CH₄ substrato FVW, B) Superfície usando banco %CH₄ refinado completo, C) Superfície usando banco %CH₄ substrato FW, D) Superfície usando banco %CH₄ refinado completo.

A análise comparativa entre as superfícies da Figura 52 mostra que as novas superfícies geradas a partir dos novos bancos de dados reforçam os resultados encontrados pelo modelo anterior, apontando para uma faixa ótima de pH para a produção de metano entre 6,5 e 7,6. Para o substrato FW é importante notar que o novo modelo aponta a obtenção de maiores concentrações volumétricas de metano para volumes de reatores maiores, enquanto a anterior apontava para volumes menores. Analisando o banco de dados é possível justificar esse comportamento, uma vez que o banco de dados %CH₄ substrato FW possui a maioria dos dados na faixa de volume de reator entre 44L e 55L, enquanto o banco de dados %CH₄ refinado completo possuía a maioria dos dados para volumes de até 19L.

A seguir é apresentada as figuras para OLR versus volume de reator.

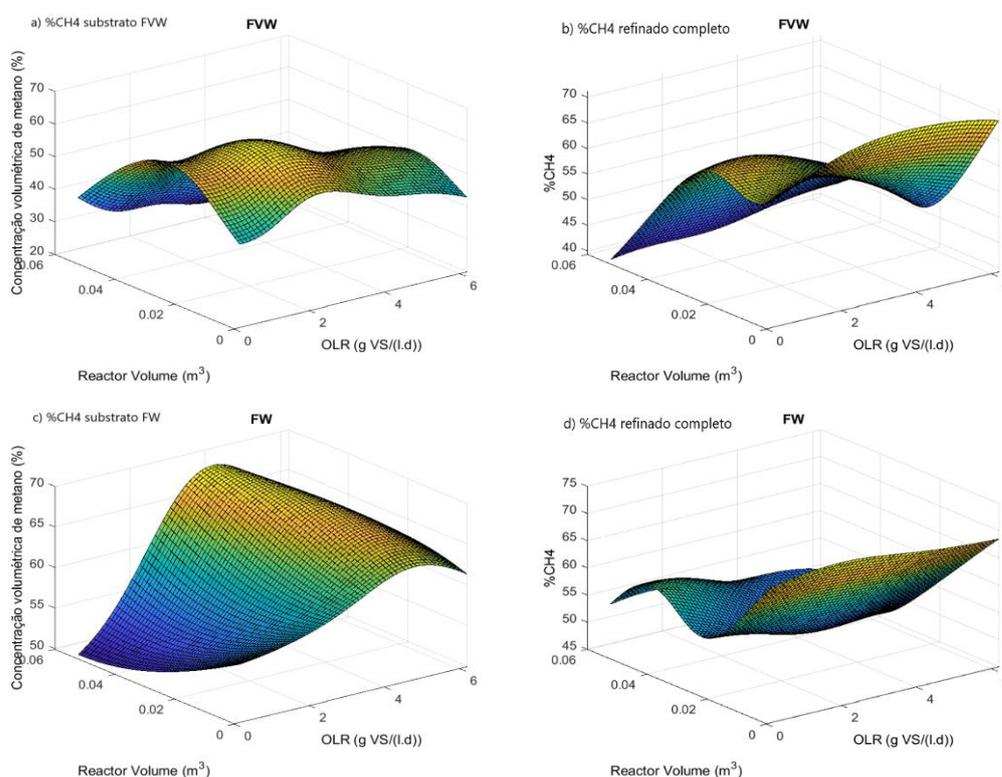


Figura 53 - Comparação entre as superfícies de resposta para a combinação entre OLR x volume de reator para %CH₄. A) Superfície usando banco %CH₄ substrato FVW, B) Superfície usando banco %CH₄ refinado completo, C) Superfície usando banco %CH₄ substrato FW, D) Superfície usando banco %CH₄ refinado completo.

A observação da Figura 53 evidencia que, para o substrato FVW, o novo modelo manteve um comportamento consistente, já previsto pelo modelo anterior. Por outro lado, para o substrato FW, houve uma notável melhoria no

comportamento da superfície, destacando claramente uma região ótima de pH entre 2 e 5 g.VS/l.d. Isso indica que o banco de dados especializado por biomassa demonstrou uma capacidade aprimorada em prever a variação da concentração volumétrica de metano com o OLR, comportamento esse que era pouco observado no modelo anterior.

A seguir é apresentada as figuras para temperatura versus volume de reator.

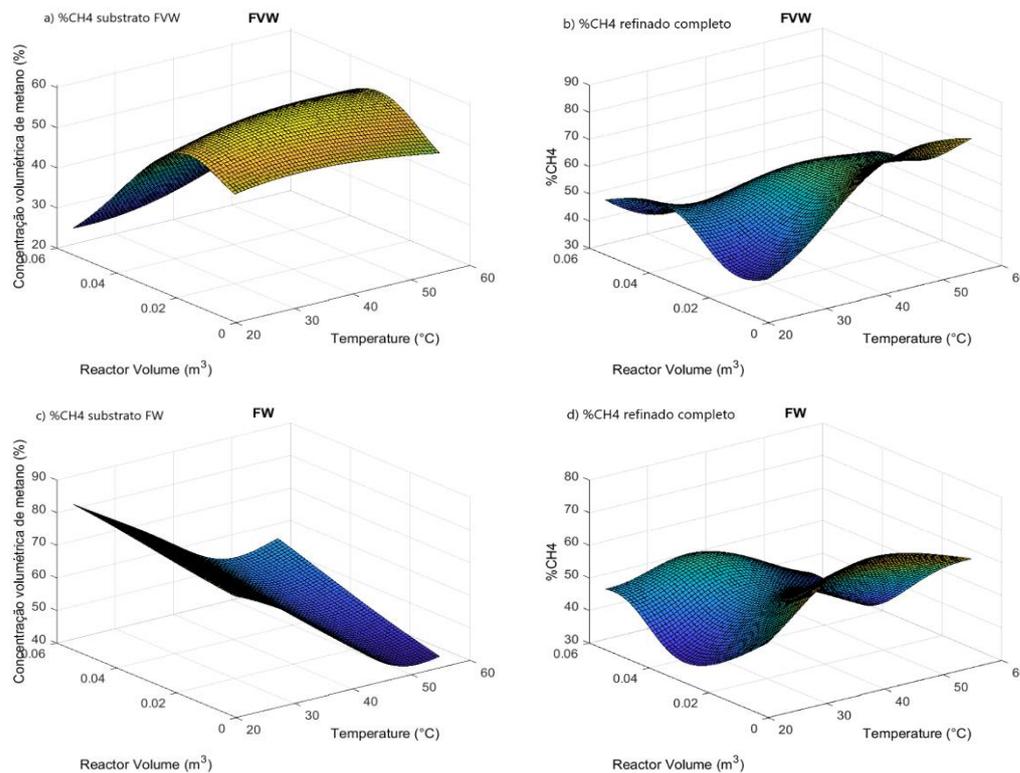


Figura 54 - Comparação entre as superfícies de resposta para a combinação entre temperatura x volume de reator para %CH₄. A) Superfície usando banco %CH₄ substrato FVW, B) Superfície usando banco %CH₄ refinado completo, C) Superfície usando banco %CH₄ substrato FW, D) Superfície usando banco %CH₄ refinado completo.

Da Figura 54 é possível observar que para o substrato FVW o novo modelo obtido a partir do banco de dados especializado (a) foi capaz de gerar uma superfície mais uniforme, ressaltando os pontos ótimos da superfície anterior (b), como concentração volumétrica de metano maior na região termofílica de temperaturas, entre 40°C e 50°C, e volume de reator menor. A superfície (c) para o substrato FW mostrou um resultado diferente do esperado pela literatura e pela seção 4.3.3, que indicam que a região termofílica de temperatura é a mais favorável para a produção de metano. Isso pode ser explicado pela falta de dados no banco de dados %CH₄

substrato FW para temperaturas inferiores a 35°C, o que pode ter afetado a precisão da rede neural e motivado a extrapolação dos resultados.

A partir dos resultados encontrados é possível dizer que a utilização de bancos de dados especializados por biomassa representa uma estratégia promissora para aprimorar a precisão dos modelos de redes neurais no contexto da estimativa da produção de biogás e metano a partir de substratos variados, como FVW e FW. Contudo, é imperativo que esses bancos de dados sejam abrangentes em relação às variáveis críticas que influenciam o processo, incluindo temperatura, volume do reator e teor de sólidos voláteis. A ausência de dados abrangentes nessas áreas pode levar a resultados inconsistentes e imprecisos nos modelos desenvolvidos. A integridade e representatividade dos dados são fundamentais para garantir a confiabilidade e a validade das previsões realizadas pelos modelos de rede neural.

4.4.4 Avaliação dos modelos de produção de biogás e metano com o conjunto de dados específico por biomassa

Foram desenvolvidos novos modelos de RNA para prever a produção de biogás e metano a partir de diferentes tipos de biomassa, utilizando bancos de dados especializados por substrato. Os novos modelos apresentaram ganhos e vantagens em relação aos modelos anteriores, como maior precisão, regularidade e identificação das condições ótimas de operação do reator fornecendo insights valiosos, destacando regiões ótimas de pH e influências cruciais da temperatura, volume do reator e VS. No entanto, os novos modelos também apresentaram limitações e desafios, como a restrição de dados e a extrapolação de valores para regiões pouco representadas no banco de dados, revelando a necessidade de dados abrangentes. Portanto, conclui-se que a disponibilidade dos dados é um fator crucial para melhorar a confiabilidade dos modelos de RNA para a produção de biogás e metano específica por biomassa.

5 Conclusão e sugestões para trabalhos futuros

Com base nos resultados obtidos neste estudo, pode-se concluir que a etapa da construção do banco de dados inédito foi fundamental para avaliar o desempenho das RNAs na previsão da produção acumulada de biogás, produção acumulada de metano e concentração volumétrica de metano nas amostras. O trabalho exaustivo de pesquisa bibliográfica para a obtenção dos 47 artigos científicos analisados permitiu a consolidação de informações em um banco de dados abrangente e inédito, contendo 2098 amostras com informações disponíveis sobre 9 variáveis do processo de biodigestão (tipo de substrato, tipo de reator, TS, VS, pH, OLR, HRT, temperatura e volume do biodigestor), além de informações sobre a produção acumulada de biogás, produção acumulada de metano e a concentração volumétrica de metano. Desse modo, esse banco de dados pode ser utilizado em estudos futuros para avaliar outras variáveis do processo na produção de biogás e metano. A testagem de diversas topologias também foi uma etapa importante deste estudo. Isso ocorre porque diferentes topologias podem afetar o desempenho das RNAs na previsão da produção de biogás e metano. Portanto, foram testadas várias topologias para identificar aquela que apresentava o melhor desempenho na previsão da produção acumulada de biogás, produção acumulada de metano e concentração volumétrica de metano nas amostras. A utilização de redes neurais artificiais (RNAs) e análise estatística mostrou-se uma abordagem eficaz para prever a produção acumulada de biogás (PCB) e metano (PCM) a partir de resíduos orgânicos, com valores significativos de R^2 e SSE. Para PAB, a melhor topologia de RNA obtida foi *trainbr_tansig_11_purelin* (algoritmo de treinamento, função de ativação da camada oculta, número de neurônios da camada oculta e função da camada de saída, respectivamente) e o valor de R^2 obtido no treinamento da RNA foi de 0,995; para PAM a melhor topologia foi *trainbr_logsig_8_purelin* com o R^2 de treinamento de 0,996 e para concentração volumétrica de metano a melhor topologia foi *trainbr_logsig_10_purelin* com R^2 de 0,901 para o treinamento. Além disso as RNAs obtiveram baixos valores de SSE para PAB, PAM e concentração volumétrica de metano (0,181, 0,109 e 11,45, respectivamente) e mantiveram a qualidade no desempenho nas etapas de teste e validação das redes. Esses resultados

indicam que as RNAs são uma ferramenta promissora para modelar processos complexos como a produção de biogás.

A combinação de diferentes variáveis do processo para a produção de superfícies de resposta é uma abordagem útil para identificar pontos ótimos no processo produtivo. Isso ocorre porque a produção de biogás é influenciada por várias variáveis, tais como tipo de substrato, tipo de reator, VS, OLR, pH, HRT, temperatura e volume de reator. Ao combinar essas variáveis em diferentes proporções, foi possível identificar as condições ideais para maximizar a produção de biogás e metano. Por exemplo, para a concentração volumétrica de metano, o ponto ótimo foi encontrado em temperaturas mesofílicas entre 30°C e 40°C e valores de pH entre 6,4 e 7,8. No entanto, é importante ressaltar que os resultados obtidos a partir das superfícies de resposta devem ser avaliados quanto à sua coerência física antes da aplicação prática desses resultados. Isso ocorre porque nem todos os resultados obtidos podem ser fisicamente coerentes ou viáveis para aplicação prática. Portanto, é importante considerar não apenas os resultados obtidos a partir das superfícies de resposta, mas também as limitações práticas do processo produtivo do biogás e metano para garantir sua aplicabilidade prática. Essas limitações dependem de diversos fatores, como o tipo de resíduo utilizado, as características do reator e as condições de operação. As limitações mais comuns incluem a necessidade constante de manutenção do reator, limitação da capacidade de processamento em relação ao volume de resíduos gerados, ocorrência de inibição do processo devido a compostos tóxicos ou falta de nutrientes, e a necessidade de controle rigoroso das condições de operação para garantir a estabilidade do reator.

Além disso, com base na meticulosa análise dos subconjuntos específicos por biomassa e na comparação detalhada dos modelos especializados com o modelo global, esta pesquisa proporcionou uma visão aprofundada das complexidades da produção de biogás e metano a partir de diferentes biomassas. As novas topologias identificadas demonstraram alta capacidade preditiva, evidenciada pelos valores de SSE e R^2 . Para o substrato FVW, os valores foram 2,2187 e 0,9741 para PAB, 1,9488 e 0,9750 para PAM, e 3,6019 e 0,9235 para %CH₄, respectivamente. No caso do substrato FW, os valores correspondentes foram 0,0029 e 0,9999 para PAB, 0,0044 e 0,9997 para PAM, e 5,3501 e 0,8367 para %CH₄, respectivamente. As variações identificadas nas propriedades de produção, especialmente em relação ao

pH, volume do reator e relação de sólidos voláteis (VS), revelaram padrões coerentes e incoerentes nos dados a partir das novas superfícies de respostas geradas. Esta compreensão refinada dos processos de produção de biogás é crucial para a otimização eficaz desses sistemas. Além disso, a metodologia adotada estabeleceu um sólido fundamento para pesquisas futuras e aplicações práticas no campo da energia renovável. Estes resultados destacam a importância da abordagem especializada na análise de dados complexos e variados como os aqui encontrados, fornecendo insights valiosos para avanços futuros nesta área.

É importante ressaltar que a universalização das RNAs para diferentes sistemas de produção de biogás requer mais pesquisas e análises para avaliar a robustez e generalização dos modelos. Ainda há muito a ser explorado nesse campo, mas os resultados do presente estudo apontam para a possibilidade de desenvolver modelos mais precisos e eficientes para prever a produção de biogás e metano a partir de resíduos orgânicos. Com o avanço da tecnologia e o acesso a um maior número de dados, a universalização das RNAs pode ser alcançada, permitindo um melhor aproveitamento de resíduos orgânicos e contribuindo para uma produção de energia mais sustentável. Em suma, este estudo contribuiu para um melhor entendimento do processo de produção de biogás e metano a partir de resíduos orgânicos e demonstrou que as RNAs e análise estatística são ferramentas promissoras para modelar esse processo.

Além disso, é relevante destacar que algumas informações, como a precisão experimental e acurácia das medições, não puderam ser recuperadas da literatura utilizada, o que pode ter afetado os resultados obtidos. Assim, sugere-se que futuras pesquisas incluam medidas experimentais mais precisas para aprimorar a qualidade dos dados. Ademais, sugestões para trabalhos futuros incluem:

- Investigar outras variáveis de processo na produção de biogás como a concentração de nutrientes (nitrogênio, fósforo e potássio), pressão no reator e a presença de contaminantes no substrato por exemplo;
- Avaliar o desempenho das RNAs em diferentes condições experimentais: É importante lembrar que o desempenho da RNA pode variar em função de diferentes parâmetros, tais como número de camadas, número de neurônios por camada, função de ativação, entre outros. Portanto, para avaliar o desempenho das RNAs em diferentes condições experimentais, é necessário

testar configurações diferentes das utilizadas no presente estudo e verificar qual delas apresenta o melhor desempenho. Além disso, é importante avaliar o desempenho das RNAs em diferentes tipos de dados experimentais, para verificar se elas são capazes de modelar com precisão diferentes conjuntos de dados;

- Comparar o desempenho das RNAs com outras técnicas de modelagem: Além das RNAs, existem outras técnicas de modelagem que podem ser usadas para analisar o processo de produção de biogás. Algumas dessas técnicas incluem a modelagem baseada em equações diferenciais, a modelagem baseada em redes de Petri e a modelagem baseada em lógica fuzzy. A modelagem baseada em equações diferenciais é adequada para sistemas complexos e pode ser usada para descrever o comportamento dinâmico de um processo. A modelagem baseada em redes de Petri é uma técnica gráfica que permite descrever visualmente a interação entre as variáveis do processo. A modelagem baseada em lógica fuzzy é uma técnica matemática que permite lidar com incertezas e imprecisões na modelagem;
- Investigar a viabilidade da aplicação da metodologia proposta em escala indústria: Investigar a viabilidade da aplicação da metodologia proposta em escala industrial é um importante aspecto para o avanço da tecnologia de produção de biogás. A escala industrial é muito diferente da escala de laboratório ou piloto, onde muitos dos estudos são realizados. Na escala industrial, existem muitos desafios a serem enfrentados, como a seleção de substratos, o projeto e operação dos reatores em grande escala, o gerenciamento e uso do biogás produzido, entre outros. A aplicação da metodologia proposta em escala industrial deve ser realizada de forma cuidadosa e planejada. É necessário avaliar aspectos como a eficiência do processo em escala maior (uma das limitações do presente estudo, que contou com poucos dados disponíveis para grandes volumes), a viabilidade econômica da produção de biogás em grande escala e a possibilidade de produzir biogás e metano de qualidade suficiente para uso em diferentes aplicações. Além disso, a aplicação em escala industrial deve ser adaptada às condições específicas de cada local, como as características do substrato

disponível, as condições climáticas e a infraestrutura disponível para a produção de biogás;

- Realizar estudos adicionais para avaliar a sustentabilidade ambiental e econômica: A produção de biogás e metano a partir de resíduos orgânicos pode ser uma opção sustentável para gerar energia renovável e reduzir resíduos. A geração distribuída de energia, que consiste em gerar energia elétrica ou térmica localmente, utilizando o biogás em pequenas usinas próximas ao local de produção de resíduos, pode trazer benefícios como a redução da dependência de fontes fósseis, a criação de empregos locais e a diminuição das emissões de gases de efeito estufa. No entanto, é necessário avaliar a eficiência do processo de produção, bem como seus impactos ambientais e econômicos, por meio de estudos que avaliem a sustentabilidade ambiental e econômica da produção de biogás e metano para a geração distribuída de energia, incluindo o impacto ambiental da utilização de substratos na produção, a qualidade do efluente gerado durante o processo e a viabilidade econômica da instalação e operação das usinas. A realização desses estudos pode ajudar a determinar se essa produção é uma opção viável e como pode ser otimizada para maximizar a eficiência e reduzir custos.

Dessa forma, este estudo abre caminho para novas pesquisas que possam contribuir para a produção mais eficiente de biogás e metano a partir de resíduos orgânicos, o que pode ter impacto significativo no desenvolvimento de fontes de energia renováveis e na geração distribuída de energia especialmente no contexto da transição para uma matriz energética mais limpa e sustentável.

6 Referências bibliográficas

- ABATZOGLOU N.; BOIVIN S. A review of biogas purification processes. *Biofuels, Bioproducts, Biorefinery*, v. 3, p. 42-71, 2009.
- AGYEMAN, F. O.; TAO, W. Anaerobic co-digestion of food waste and dairy manure: effects of food waste particle size and organic loading rate. *Journal of Environmental Management*, v. 133, p. 268-274, 2014.
- AKBAS, H.; BILGEN, B.; TURHAN, A.M. An integrated prediction and optimization model of biogas production system at a wastewater treatment facility. *Bioresource Technol*, v. 196, p. 566–76, 2015.
- ALI, A. M.; ALAM, M. Z.; ABDUL-LATIF, F. M.; JAMI, M. S.; BOUH, I. G.; BELLO, I. A.; AINANE, T. Production of biogas from food waste using the anaerobic digestion process with biofilm-based pretreatment. *Processes*, v. 11, n. 3, 2023.
- AMARAL, A. C.; STEINMETZ, R. L. R.; KUNZ, A. Capítulo I - Processo de Biodigestão. In: KUNZ, A.; STEINMETZ, R. L. R.; AMARAL, A. C. (Ed.). *Biodigestão Anaeróbia: Fundamentos e Aplicações*. Brasília, DF: Embrapa, 2020. p. 13-36.
- ANDERSON, D.R.; SWEENEY, D.J.; WILLIAMS, T.A.; CAMM, J.D.; COCHRAN, J.J. *Estatística aplicada à administração e economia*. 2. ed. São Paulo: Cengage Learning, 2011.
- ARAÚJO, A. P. C. Produção de biogás a partir de resíduos orgânicos utilizando biodigestor anaeróbio de leito fixo com manta de lã de vidro como suporte para imobilização de biomassa microbiana. 2017. 64 f. Monografia (Graduação em Engenharia Química) - Universidade Federal de Uberlândia, Uberlândia, 2017.
- ARAÚJO, Monique Gomes de. Previsão da curva de produção para projeto exploratório utilizando redes neurais artificiais. 2020. 88 f. Dissertação (Mestrado em Engenharia Mecânica) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2020.
- APPELS, L.; VAN ASSCHE, A.; WILLEMS, K.; DEGRÈVE, J.; VAN IMPE, J.; DEWIL, R. Peracetic acid oxidation as an alternative pre-treatment for the anaerobic digestion of waste activated. *Bioresource Technology*, v. 102, p. 4124–30, 2011.

- ALMEIDA, C. Potencial de produção de biogás a partir da biomassa de suinocultura com culturas energéticas. Dissertação (Mestrado em Engenharia Agrícola) - Universidade Estadual do Oeste do Paraná, Cascavel, 2016.
- ARUN, J.; GOPINATH, K. P.; SIVARAMAKRISHNAN, R.; MADHAV, N. V.; ABHISHEK, K.; RAMANAN, V. G. K.; PUGAZHENDHI, A. Bioenergy perspectives of cattails biomass cultivated from municipal wastewater via hydrothermal liquefaction and hydro-deoxygenation. *Fuel*, v. 284:1-27, 2021.
- ASSOCIAÇÃO BRASILEIRA DO BIOGÁS – ABIOGÁS. Biogás no Brasil: conhecendo o mercado no país [2021]. Disponível em: <https://abiogas.org.br/wpcontent/uploads/2021/01/InfograficosAbiogas_D_2021-1.pdf>. Acesso em: 9 nov. 2022.
- BARIK, D.; MURUGAN, S. An artificial neural network and genetic algorithm optimized model for biogas production from co-digestion of seed cake of Karanja and cattle dung. *Waste Biomass Valor*, v. 6, p. 1015–27, 2015.
- BASHEER, I.A.; HAJAMMAHMOOD, M. Artificial neurais networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, v. 43, n. 1, p. 3-31, 2000.
- BEUX, S. Avaliação do tratamento de efluente de abatedouro em digestores anaeróbios de duas fases. Dissertação (Mestrado em Engenharia Química) - Universidade Estadual de Ponta Grossa, Ponta Grossa, 2005.
- BEHERA, S. K.; MEHER, S. K.; PARK, H.-S. Artificial neural network model for predicting methane percentage in biogas recovered from a landfill upon injection of liquid organic waste. *Clean Technologies and Environmental Policy*, v. 17, p. 443-53, 2015.
- BELTRAMO, T.; KLOCKE, M.; HITZMANN, B. Prediction of the Biogas Production Using GA and ACO Input Features Selection Method for ANN Model. *Information Processing in Agriculture*, v. 6, n. 3, p. 349-56, 2019.
- Biogás — Conheça o Potencial e Importância Dessa Fonte de Energia. Ecogen Brasil, 2021. Disponível em: <<https://ecogenbrasil.com.br/blog/biogas-conheca-o-potencial-e-importancia-dessa-fonte-de-energia>>. Acesso em: 30 out. 2022.
- BRAGA ET AL., A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. *Redes Neurais Artificiais – Teoria e Aplicações*. Rio de Janeiro: LTC, 2000.

- BRASIL. Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. Resolução ANP nº 8, de 30 de janeiro de 2015. Estabelece a especificação do Biometano contida no Regulamento Técnico ANP nº 1/2015, parte integrante desta Resolução. Diário Oficial da União, Brasília, DF, 2 fev. 2015a. Seção 1, p. 29-31.
- BRASIL. Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. Resolução ANP nº 8, de 30 de janeiro de 2015. Estabelece a especificação do Biometano contida no Regulamento Técnico ANP nº 1/2015, parte integrante desta Resolução. Disponível em: <<https://www.legisweb.com.br/legislacao/?id=280722>>. Acesso em: 30 out. 2022.
- BRASIL. Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. Resolução ANP nº 685, de 29 de junho de 2017. Estabelece as regras para aprovação do controle da qualidade e a especificação do biometano oriundo de aterros sanitários e de estações de tratamento de esgoto destinado ao uso veicular e às instalações residenciais. Diário Oficial da União, Brasília, DF, 30 jun. 2017. Seção 1, p. 64-66. Disponível em: <<https://www.legisweb.com.br/legislacao/?id=345545>>. Acesso em: 30 out. 2022.
- CARLSSON, M.; LAGERKVIST, A.; ECKE H. Electroporation for enhanced Methane yield from municipal solid waste. In: MOVING ORGANIC WASTE RECYCLING TOWARDS RESOURCE MANAGEMENT AND BIOBASED ECONOMY (ORBIT '08), 6., 2008, Wageningen. Proceedings... Wageningen: ORBIT Association, 2008. p. 1-8.
- CARVALHO, A. L.; BORBA, B. M.; SANTOS, E. M.; CORTES, M. R.; PEREIRA, E. P.; SANTOS, M. T. Biogás e biometano: potencialidades e desafios para o Brasil. Revista Brasileira de Energia, v. 25, n. 1, p. 9-28, 2019.
- CAMARGO, D.W.N. Uso de residuos agrícolas para la producción de biocombustibles em el departamento del Meta. Tecnura, v. 16, p. 142-156, 2012.
- CECCON, D. Funções de ativação: definição, características, e quando usar cada uma [2020]. Disponível em: <<https://iaexpert.academy/2020/05/25/funcoes-de-ativacao-definicao-caracteristicas-e-quando-usar-cada-uma/>>. Acesso em março 2022.

- CIBIOGÁS. Nota Técnica: N° 002/2010 – Panorama do Biogás no Brasil em 2019. Foz do Iguaçu, abril de 2020.
- CORTEZ, L. A. B.; LORA, E. E. S.; GÓMEZ, E. O. Biomassa para energia. Campinas: Editora da Unicamp, 2008.
- COELHO, S. T.; VELÁZQUEZ, S. M. S. G.; SILVA, O. C.; PECORA, V.; ABREU, F. C. Geração de energia elétrica a partir do biogás proveniente do tratamento de esgoto. In: CONGRESSO BRASILEIRO DE ENERGIA, 11., 2006, Rio de Janeiro. Anais... Rio de Janeiro: PIPGE, 2006. p. 1-5.
- CHEN, Y.; CHAN, Y. J.; LIM, J. W.; LIEW, C. S.; MOHAMAD, M.; HO, C. D.; USMAN, A.; LISAK, G.; HARA, H.; TAN, W. N. Artificial Neural Network (ANN) Modelling for Biogas Production in Pre-Commercialized Integrated Anaerobic-Aerobic Bioreactors (IAAB). *Water*, v. 14, n. 9, 2022.
- CHEN, W.-Y.; CHAN, Y. J.; LIM, J. W.; LIEW, C. S.; MOHAMAD, M.; HO, C.-D.; USMAN, A.; LISAK, G.; HARA, H.; TAN, W.-N. Artificial Neural Network (ANN) Modelling for Biogas Production in Pre-Commercialized Integrated Anaerobic-Aerobic Bioreactors (IAAB). *Water*, v. 14, n. 9, p. 1410, 2022.
- CHERNICHARO, C.A.L. Princípios do tratamento biológico de águas residuárias; reatores anaeróbios. Belo Horizonte: Departamento de Engenharia Sanitária e Ambiental, UFMG, 1997. v. 5.
- CHU, C.-F.; LI, Y.-Y.; XU, K.-Q.; EBIE, Y.; INAMORI, Y.; KONG, H.-N. A pH- and temperature-phased two-stage process for hydrogen and methane production from food waste. *International Journal of Hydrogen Energy*, v. 33, n. 18, p. 4739–4746, 2008.
- CYBENKO, G. Approximation by superpositions of a sigmoid funeration process. *Applied Artificial Intelligence*, v. 11, n. 1, p. 33-57, 1997.
- DAS GHATAK, M.; GHATAK, A. Artificial neural network model to predict behavior of biogas production curve from mixed lignocellulosic co-substrates. *Fuel*, v. 2018.
- DEMIREL B.; SCHERER P.A... The roles of acetotrophic and hydrogenotrophic methanogens during anaerobic conversion of biomass to methane: a review. *Reviews in Environmental Science and Bio/Technology*, v7, n3, p173-190 ,2008

- EL-MASHAD, H.M.; ZHANG, R... Biogas production from co-digestion of dairy manure and food waste. *Bioresource Technology*, v 101, n 11, p 4021-4028.,2010.
- FIDALGO, J.P.G. Modelagem usando inteligência artificial para estudar o pré tratamento de biomassa lignocelulósica. [Dissertação]. Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro ,2019.
- FNR - Fachagentur Nachwachsende Rohstoffe. Guia Prático do Biogás: Geração e Utilização. Ministério da Nutrição Agricultura e Defesa do Consumidor da Alemanha. Gülzow, Alemanha ,2010
- FORSTER-CARNEIRO, T.; ROMERO, M.; ROMERO, L. I. Influence of total solid and inoculum contents on performance of anaerobic reactors treating food waste. *Bioresource Technology*, v. 99, n. 15, p. 6994–7002, 2008.
- GALBIATTI, J. A.; CAMELO, A. D.; SILVA, F. G.; GERARDI, E. A. B.; CHICONATO, D. A. Estudo qualitativo do biogás produzido por substratos em biodigestores tipo batelada. *Revista Brasileira de Engenharia Agrícola e Ambiental*, Campina Grande, v. 14, n. 4, p. 432-437, 2010.
- GALLERT, C.; BAUER, S.; WINTER, J. Effect of ammonia on the anaerobic degradation of protein by mesophilic and thermophilic biowaste. *Applied Microbiology and Biotechnology*, Berlin, v. 50, n. 4, p. 495-501, oct. 1998.
- GALLERT, C.; WINTER, J. Mesophilic and thermophilic anaerobic digestion of source-sorted organic wastes: effect of ammonia on glucose degradation and methane production. *Applied Microbiology and Biotechnology*, Berlin, v. 48, n. 3, p. 404-410, sep. 1997.
- GEHRING, C. G. Análise da geração de energia elétrica a partir do biogás produzido na fermentação anaeróbica de vinhaça. 2014. 124 p. Trabalho de conclusão de curso (Engenharia Elétrica) – Escola de Engenharia Elétrica de São Carlos, Universidade de São Paulo, São Carlos – SP.
- GHATAK, M. D.; GHATAK, A. Artificial Neural Network Model to Predict Behavior of Biogas Production Curve from Mixed Lignocellulosic Co-Substrates. *Fuel*, v. 232, p. 178-89, 2018.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep learning. Cambridge: MIT press, 2016.

- GUEGUIM KANA, E.; OLOKE, J.; LATEEF, A.; ADEWALE, A. Modeling and optimization of biogas production on saw dust and other co-substrates using Artificial Neural Network and Genetic Algorithm. *Renewable Energy*, v. 44, p. 215-220, 2012.
- GONÇALVES NETO, J.; VIDAL OZORIO, L.; CAMPOS DE ABREU. C.; FERREIRA DOS SANTOS, B.; ALAIN YANNICK PRADELLE, F. Modelling of biogas production from food, fruits and vegetables wastes using artificial neural network (ANN). *Fuel journal*, v. 2020.
- GUIMARÃES, C.; MAIA, D.; SERRA, E. Construction of Biodigesters to Optimize the Production of Biogas from Anaerobic Co-Digestion of Food Waste and Sewage. *Energies*, v. 11, n. 4, p. 870, 9 abr. 2018.
- HAYKIN, S. *Redes Neurais- Princípios e Práticas*. Porto Alegre: Bookman, 2001.
- HAGEN, L. H.; MÜLLER, B.; RUSTEN, B.; LINJORDET, R.; LIU, J.; KANTZOW, C.; HORN, S. J. Microbial community structure and dynamics during co-digestion of whey permeate and cow manure in continuous stirred tank reactor systems. *Bioresource Technology*, v. 171, p. 350-359, 2014.
- HU, J.; ZHANG, Y.; WANG, Y.; ZHANG, L.; HE, P.; LI, G. Enhanced anaerobic digestion of vegetable waste by adding high concentrations of fat oil and grease. *Waste Management*, v. 91, p. 1-8, 2019.
- INSTITUTO 17. *Biogás no Brasil: Potencial Oferta a Curto Prazo*. Programa de Energia para o Brasil – BEP (Brasil). Relatório técnico 02-2021. São Paulo: Instituto 17, 2021.
- IPCC. *IPCC Second Assessment Report: Climate Change 1995 (SAR) - Working Group I: The Science of Climate Change*. Cambridge: IPCC, 1996
- IRENA. *Perspectiva da Transição Energética Mundial Caminho dos 1, 5°C*. Abu Dhabi: IRENA, 2021. Disponível em: <https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2021/Jun/IRENA_WETO_Summary_2021_PT.pdf>. Acesso em: 30 out 2022.
- JAROENPOJ, S.; YU, Q. J.; NESS, J. Development of Artificial Neural Network Models for Biogas Production from Co-Digestion of Leachate and Pineapple Peel. *The Global Environmental Engineers*, v. 1, p. 42-7, 2014.

- JIANG, J.; ZHANG, Y.; LI, K.; WANG, Q.; GONG, C.; LI, M. Volatile fatty acids production from food waste: Effects of pH, temperature, and organic loading rate. *Bioresource Technology*, v. 143, p. 525–530, 2013.
- I, C.; KONG, C.-X.; MEI, Z.-L.; LI, J. A Review of the Anaerobic Digestion of Fruit and Vegetable Waste. *Applied Biochemistry and Biotechnology*, v. 183, n. 3, p. 906–922, 2017.
- KARDAM, A.; RAJ, K. R.; ARORA, J. K.; SRIVASTAVA, S. Artificial neural network modeling for biosorption of Pb (II) ions on nanocellulose fibers. *BioNanoScience*, v. 2, p. 153–160, 2012.
- KARLSSON, T. Manual básico de biogás. Lajeado: Editora UNIVATES, 2014.
- KETHA, S. Efeito de outliers no desempenho de redes neurais. *Medium*, 2019.
- KHALID, A.; ARSHAD, M.; ANJUM, M.; MAHMOOD, T.; DAWSON, L. The anaerobic digestion of solid organic waste. *Waste Management*, v. 31, n. 8, p. 1737-1744, 2011.
- KHALIL, A.; ALMASRI, M. N.; MC KEE, M.; KALUARACHCHI, J.J. Applicability of statistical learning algorithms in groundwater quality modeling. *Water Resources Research*, v. 41, p. 1-16, 2005.
- KHADKA, A.; PARAJULI, A.; DANGOL, S.; THAPA, B.; SAPKOTA, L.; CARMONA-MARTÍNEZ, A. A.; GHIMIRE, A. Effect of the substrate to inoculum ratios on the kinetics of biogas production during the mesophilic anaerobic digestion of food waste. *Energies*, v. 15, n. 3, p. 1-18, 2022.
- KOMEMOTO, K.; LIM, Y. G.; NAGAO, N.; ONOUE, Y.; NIWA, C.; TODA, T. Effect of temperature on VFA's and biogas production in anaerobic solubilization of food waste. *Waste Management*, v. 29, n. 6, p. 2950–5, 2009.
- KUNZ, A.; STEINMETZ, R. L. R.; DO AMARAL, A. C. Fundamentos da digestão anaeróbia, purificação do biogás, uso e tratamento do digestato. Embrapa Suínos e Aves-Livro científico (ALICE), 2019.
- KUNWAR, P.; KUSHWAHA, S. K.; YADAV, M.; PAREEK, N.; CHAWADE, A.; VIVEKANAND, V. Food Waste to Energy: An Overview of Sustainable Approaches for Food Waste Management and Nutrient Recycling. *BioMed Research International*, v. 2017, p. 1-19, 2017.

- KOIKE, Y.; MORIMURA, S.; WU, X.-L.; KIDA, K. Production of fuel ethanol and methane from garbage by high-efficiency two-stage fermentation process. *Journal of Bioscience and Bioengineering*, v. 108, n. 6, p. 508-512, 2009.
- LEE, D.-H.; KIM, Y.-T.; LEE, S.-R. Shallow Landslide Susceptibility Models Based on Artificial Neural Networks Considering the Factor Selection Method and Various Non-Linear Activation Functions. *Remote Sens*, v. 12, n. 7, p. 1194, 2020.
- LEE, J. P.; LEE, J. S.; PARK, S. C. Two-phase methanization of food wastes in pilot scale. *Applied Biochemistry and Biotechnology*, v. 77-v. 79, p. 585-593, 1999.
- LIU, Y.; ZHANG, Y.; LI, K.; WANG, Q.; GONG, C.; LI, M. Influence of pH on anaerobic co-digestion of kitchen waste and waste activated sludge in a pilot-scale anaerobic digestion system. *Bioresource Technology*, v. 275, p. 324-331, 2019.
- LIM, S.-J.; KIM, B. J.; JEONG, C. M.; CHOI, J. D.; AHN, Y. H.; CHANG, H. N. Anaerobic organic acid production of food waste in once-a-day feeding and drawing-off bioreactor. *Bioresource Technology*, v. 99, p. 7866–7874, 2008.
- LUDWIG, O.; COSTA, E.M.M. *Redes Neurais: Fundamentos e Aplicações com Programas em C*. Rio de Janeiro: Editora Ciência Moderna, 2007.
- MAHANTY, B.; ZAFAR, M.; PARK, H.-S... Characterization of co-digestion of industrial sludges for biogas production by artificial neural network and statistical regression models. *Environmental Technology*, v. 34, n. 13-14, p. 2145-2153., 2013.
- MATA-ALVAREZ, J.; MACÉ, S.; LLABRÉS, P. Anaerobic digestion of organic solid wastes: An overview of research achievements and perspectives. *Bioresource Technology*, v. 160, p. 692-700, 2014.
- MATOS, C.F.; PAES, J.L.; PINHEIRO, É.F.; DE CAMPOS, D.V. Biogas production from dairy cattle manure, under organic and conventional production systems. *Engenharia Agrícola*, v. 37, n. 6, p. 1081-1090, 2017.
- MATHEUS. *Funções de Ativação: Entendendo a importância da ativação correta nas redes neurais*. Disponível em: <<https://matheusfacure.github.io/2017/07/12/activ-func/>>. Acesso em março 2022.

- MATEUS, M. S. C. S.; AGUIAR, P. H. V.; BARROS, R. T. V.; OLIVEIRA, E. C.; GADÉA, A. M. Estimativa da geração de biogás a partir de RSU utilizando redes neurais artificiais (RNA). 30º Congresso ABES, 2019.
- MASSANET-NICOLAU, J.; DINSDALE, R.; GUWY, A.; SHIPLEY, G. Use of real time gas production data for more accurate comparison of continuous single-stage and two-stage fermentation. *Bioresource Technology*, vol. 129, pp. 561–567, 2013.
- MICROSOFT. Tipos de dados (Transact-SQL) -SQL Server. Microsoft Learn ,2023.
- MINISTÉRIO DAS CIDADES. Secretaria Nacional de Saneamento Ambiental. Conceitos para o licenciamento ambiental de usinas de biogás / Probiogás. Brasília, 2016. Disponível em: <https://antigo.mdr.gov.br/images/stories/ArquivosSNSA/probiogas/licenciamento-usinas-biogas.pdf>. Acesso em: 13 jan. 2023.
- MILANEZ, A. Y.; GUIMARAES, D. D.; MAIA, G. B. S.; SOUZA, J. A. P.; LEMOS, M. L. F. Biogás de resíduos agroindustriais: panorama e perspectivas. *BNDES Setorial*, Rio de Janeiro, n. 47, p. 221-275, 2018.
- MORAIS, G. Bioenergia: Qual sua importância e por que adotá-la? TOTVS ,11 ago2021. Disponível em: <<https://www.totvs.com/blog/gestao-agricola/bioenergia/>>. Acesso em: 30 out2022.
- MOUGARI, F; BENYAHIA, B; BELHADJ, M. Artificial neural network modeling of biogas production from anaerobic co-digestion of organic wastes. *Renewable Energy*., v 164, p 1482-1491., 2021.
- NATIVIDAD PEREZ-CAMACHO M.; CURRY R.; CROMIE T. Life cycle environmental impacts of biogas production and utilisation substituting for grid electricity, natural gas grid and transport fuels. *Waste Manage* 95:90–101. 2019.
- NAGAO, N., TAJIMA, N., KAWAI, M. et al., “Maximum organic loading rate for the single-stage wet anaerobic digestion of food waste,” *Bioresource Technology*, vol. 118, pp. 210–218, 2012.
- OLIVEIRA, A. J.; RAMALHO, J. Plano Nacional de Agroenergia 2006-2011 - Ministério da Agricultura, Pecuária e Abastecimento, Secretaria de Produção e Agroenergia. 2. ed. rev. - Brasília, DF: Embrapa Informação Tecnológica, 110 p., 2006.

- OLIVEIRA, P.A.V. (Coord.) Tecnologias para o manejo de resíduos na produção de suínos - Manual de boas práticas. 01. ed. Concórdia - SC: Embrapa, 2004. v. 01. 109 p. OLIVEIRA, P. A. V.; HIGARASHI, M. M. Geração e utilização de biogás em unidades de produção de suínos. Concórdia: Embrapa Suínos e Aves, 2006, 42 p.
- PALANISWAMY, D.; RAMESH, G.; SIVASANKARAN, S.; KATHIRAVAN, N. Optimising biogas from food waste using a neural network model. *Ice*. 2016.
- PARITOSH, K.; KUSHWAHA, S. K.; YADAV, M.; PAREEK, N.; CHAWADE, A.; VIVEKANAND, V. Food Waste to Energy: An Overview of Sustainable Approaches for Food Waste Management and Nutrient Recycling. *BioMed Research International*, v. 2017, p. 1-19, 2017.
- PARK, Y-S; LEK, S... Artificial Neural Networks. In *Developments in Environmental Modelling*. Amsterdam: Elsevier BV, 2016. p123-140.
- PORTAL DO BIOGAS. Biodigestão Anaeróbia - As 4 fases. Disponível em: <<https://portaldobiogas.com/biodigestao-anaerobia/>>. Acesso em: 30 out. 2022.
- PHUN CHIEN BONG, C.; YEE LIM, L.; TIN LEE, C.; JAROMIR FLEMES, J.; SIONG HO, C.; SHIN HO, W. The characterisation and treatment of food waste for improvement of biogás production during anaerobic digestion. A review. *Journal of Cleaner Production*. 2018.
- QDAIS, H; BANI HANI, K; SHATNAWI, N... Modeling and optimization of biogas production from a waste digester using artificial neural network and genetic algorithm. *Resources ,Conservation and Recycling*, v54, n6, p359-363., 2010.
- RAJPUT, A. A; SHEIKH, Z. Effect of inoculum type and organic loading on biogas production of sunflower meal and wheat straw. *Sustainable Environment Research*, v29, n4, p1-9., 2019.
- RAMACHANDRAN, A; RUSTUM, R; ADELOYE, A.J. Review of Anaerobic Digestion Modeling and Optimization Using Nature-Inspired Techniques Processes., v7, n12, p953., 2019.
- REN21. Market Developments | Bioenergy. 2023. Disponível em: <https://www.ren21.net/gsr-2023/modules/energy_supply/02_market_developments/01_bioenergy/>. Acesso em: 30 out 2022.

- REGIS, R. O Biogás como fonte despachável, descentralizada, descarbonizada e competitiva. [Apresentação]. VI Fórum do biogás ,2019.
- REGO, A.S.C; LEITE, S.A.F; LEITE, B.S; GRILLO, A.V; SANOTS, B.F. Artificial Neural Network Modelling for Biogas Production in Biodigesters. *Chemical Engineering Transactions*, v 74, p25-30, 2019.
- ROS, M.; FRANKE-WHITTLE, I.H.; MORALES, A.B.; INSAM, H.; AYUSO, M.; PASCUAL, J.A. Archaeal community dynamics and abiotic characteristics in a mesophilic anaerobic co-digestion process treating fruit and vegetable processing waste sludge with chopped fresh artichoke waste. *Bioresource Technology*, v. 136, p. 1-7, 2013.
- SAMPAIO, S. R.; OLIVEIRA, C. R.; GONÇALVES, R. F. Panorama da digestão anaeróbia de resíduos orgânicos com alto teor de sólidos: tendências e potencialidades. *Revista DAE*. São Paulo, 2021, p. 85-97. Disponível em: http://revistadae.com.br/artigos/artigo_edicao_231_n_1977.pdf. Acesso em: 13 jan. 2023.
- SANTOS, C. A.; FRIGO, E. P.; NOGUEIRA, C. E. C. Estudo do panorama da geração de energia a partir da biomassa de resíduos na região sul do Brasil. *Acta Iguazu*, v. 3, n. 1, p. 87-96, 2014.
- SANTOS, J. M. C. F. S. Avaliação da biomassa de plantas bioenergéticas. 2009. 79 f. Dissertação (Mestrado em Engenharia do Ambiente) - Instituto Superior de Agronomia, Universidade Técnica de Lisboa, Lisboa, 2009.
- SÁNCHEZ, E.; BORJA, R.; WEILAND, P.; TRAVIESO, L.; MARTÍN, A. Effect of substrate concentration and temperature on the anaerobic digestion of piggery waste in a tropical climate. *Process Biochemistry*, v. 37, n. 5, p. 483–9, 2001.
- SATHISH, S.; VIVEKANANDAN, S. Parametric optimization for floating drum anaerobic bio-digester using Response Surface Methodology and Artificial Neural Network. *Alexandria Engineering Journal*, v. 55, n. 4, p. 3297–3307, 2016.
- SERRANO-RUIZ, J.C. Biomass: a renewable source of fuels, chemicals and carbon materials. *Molecules*, v. 25, n. 1, p. 1-3, 2020.
- SEIXAS, J.; FOLLE, S.; MARCHETTI, D. Construção e funcionamento de biodigestores. [Manual]. Brasília: EMBRAPA-DID, Circular técnica nº 4, 1980.

- SILVA, C.A.B.V. Limpeza e Purificação de Biogás. [Dissertação]. Vila Real: Universidade de Trás-os-Montes e Alto Douro ,2009.
- SILVA, C.O.; KONRAD, O.; CALLADO, N.H.; MARDER, M.; ARAUJO, L.G. S... Resíduos sólidos orgânicos domésticos como substrato potencial para produção de biogás. *Revista Ibero Americana de Ciências Ambientais*, v11, n2, p204-212., 2020.
- SILVA, I.N.; SPATTI, D.H.; FLAUZINO, R.A. Redes Neurais Artificiais para Engenharia e Ciências Aplicadas curso prático. *Artliber* ,2010.
- SMITH, D. B.; ALMQUIST, C. B. The anaerobic co-digestion of fruit and vegetable waste and horse manure mixtures in a bench-scale, two-phase anaerobic digestion system. *Environmental Technology*, v. 35, p. 859-867, 2014.
- SOUZA, S.N.M; PEREIRA, W.C; PAVA, A.A.C... Custo da eletricidade gerada em conjunto motor gerador utilizando biogás da suinocultura. *Acta Scientiarum Technology*, v26, n2, p127-133., 2004.
- SOUSA, A. L; RIZZATTO, M.L. Produção de biogás a partir de resíduos orgânicos: uma revisão. *Scientific Electronic Archives*, v15, n2, p1-11., 2022.
- SOUZA, C. F; LUCAS, J. J; FERREIRA, W.P.M. Biodigestão anaeróbia de dejetos de suínos sob efeito de três temperaturas e dois níveis de agitação do substrato: Considerações sobre a partida. *Engenharia Agrícola*, v25, n2, p530-539., 2005.
- SOSNOWSKI, P; WIECZOREK, A; LEDAKOWICZ, S; OLESZKIEWICZ, J.A. Effect of pH control strategy on the efficiency of anaerobic digestion of dairy wastewater. *Journal of Environmental Management.*, v223, p707-713, 2018.
- STAUDT, T.B. Predição do volume de biogás produzido em sistemas de biodigestão utilizando redes neurais artificiais. [Dissertação]. Caxias do Sul: Universidade de Caxias do Sul ,2019.
- TAMPIO, E.; ERVASTI, S.; PAAVOLA, T.; HEAVEN, S.; BANKS, C.; RINTALA, J. Anaerobic digestion of autoclaved and untreated food waste. *Waste Management*, v. 34, n. 2, p. 370–377, 2014.
- TAIGANIDES E.P. Biomass-energy recovery from animal waste Part I. *World Animal Review*, v35, p2 ,1980
- TCHAKOUTÉ, H. K.; FOTIO, D.; NKENG, G. E.; JANNOT, Y. Optimization of biogas production from anaerobic digestion of water hyacinth using response

- surface methodology. *Journal of Environmental Management*, v. 222, p. 29-38, 2018.
- TSHEMESE, Z.; DEENADAYALU, N.; LINGANISO, L.Z.; CHETTY, M. An Overview of Biogas Production from Anaerobic Digestion and the Possibility of Using Sugarcane Wastewater and Municipal Solid Waste in a South African Context. *Applied System Innovation*, v. 6, n. 1, p. 13, 2023.
- THEURETZBACHER, F.; BAUER, A.; LIZASOAIN, J.; BECKER, M.; ROSENAU, T.; POTTHAST, A.; FRIEDL, A.; PIRINGER, G.; GRONAUER, A. Potential of different *Sorghum bicolor* (L. moench) varieties for combined ethanol and biogas production in the Pannonian climate of Austria. *Energy*, v. 55, p. 107-113, 2013.
- TÚPAC VALDIVIA, Y.A.; FERNANDES FILHO, J.; SILVA NETO, A.J.; SILVA, M.A.G.; SOUZA NETO, J.D. Redes neurais artificiais aplicadas à engenharia de reservatórios de petróleo: uma revisão bibliográfica. *Revista Brasileira de Geofísica*, v. 23, n. 4, p. 393-407, 2005.
- UCHIDA, M.Y. Geração Distribuída como Promotor da Política Nacional dos Resíduos Sólidos. [Apresentação]. *Abiogás*, 2019.
- VAN HAANDEL, A.C; LETTINGA, G. Tratamento Anaeróbico de Esgotos: Um Manual para Regiões de Clima Quente. [Manual]. Campina Grande: Epgraf, 1994.
- VASCO-CORREA, J; KHANAL, S; MANANDHAR, A; SHAH, A. Anaerobic digestion for bioenergy production: global status, environmental and Technoeconomic implications, and government policies. *Bioresource Technology*, v247, p1015-1026, 2018.
- WANG, X; SHI, Y; XUE, Y; WU, S; QIAO, W. Prediction of biogas production from various agricultural wastes using multi-layer perceptron neural network. *Journal of Cleaner Production*, v279, p123610, 2021.
- WANNASEK, L; ORTNER, M; KAUL, H.P; AMON, B; AMON, T. Double-cropping systems based on rye, maize and sorghum: impact of variety and harvesting time on biomass and biogas yield. *European Journal of Agronomy*, v110, p1-10, 2019.

- WEI, R.R; CHENG, G.W; LUO, J.J; ET AL. Biogas and bio-energy production from anaerobic digestion of piggery manure at different temperatures. *J Agric Mech Res*, v4, p170–4., 2010. (in Chinese).
- WIT, M; FAAIJ, A. European biomass resource potential and costs. *Biomass And Bioenergy*, v34, p188-202, 2010.
- WITEK-KROWIAK, A; CHOJNACKA, K; PODSTAWCZYK, D; DAWIEC, A; POKOMEDA, K. Application of response surface methodology and artificial neural network methods in modelling and optimization of biosorption process. *Bioresource technology*, v160, p150–160, 2014.
- YETILMEZSOY, K; TURKDOGAN, F.I; TEMIZEL, I; GUNAY, A. Development of Ann-Based Models to Predict Biogas and Methane Productions in Anaerobic Treatment of Molasses Wastewater. *International Journal of Green Energy*, v10, p885–907, 2013.
- ZANETTE, A.L. Potencial de aproveitamento energético do biogás no Brasil. [Dissertação]. Rio de Janeiro: Universidade Federal do Rio de Janeiro ,2009.
- ZARKADAS, E.; PEBAY-PEYROULA, E.; THOMPSON, M.J.; SCHOEHN, G.; UCHANSKI, T.; STEYAERT, J.; CHIPOT, C.; DEHEZ, F.; BAENZIGER, J.E.; NURY, H. Conformational transitions and ligand-binding to a muscle-type nicotinic acetylcholine receptor. *Neuron*, v. 110, n. 8, p. 1358-1370, 2022.
- ZHANG, Y.; ZHANG, J.; LIU, X.; ZHANG, C.; ZHANG, R.; LI, Y. Prediction of biogas production from anaerobic digestion using a hybrid machine learning method. *Renewable Energy*, Oxford, v. 151, p. 151-159, feb. 2020.
- ZHONG, W.; ZHANG, Z.; LUO, Y.; SUN, S.; QIAO, W.; XIAO, M. Effect of biological pretreatments in enhancing corn straw biogas production. *Bioresource Technology*, Oxford, v. 102, n. 24, p. 11177-11182, dec. 2011.
- ZHONG, L.; ZHANG, J.; CHEN, M.; SUI, Q.; WANG, R.; TONG, J.; WEI, Y. Anaerobic co-digestion of municipal sludge and food waste in a pilot-scale reactor: performance and microbial community analysis. *Environmental Technology*, Abingdon, v. 40, n. 15, p. 1919-1928, aug. 2019.