



**Fernanda da Silva Pimentel**

**Desenvolvimento de modelos preditivos para a  
gaseificação de biomassa usando redes  
neurais artificiais**

**Dissertação de Mestrado**

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

Orientador: Prof. Florian Alain Yannick Pradelle  
Coorientador: Prof. Brunno Ferreira dos Santos

Rio de Janeiro  
Maio de 2022



**Fernanda da Silva Pimentel**

**Desenvolvimento de modelos preditivos para a  
gaseificação de biomassa usando redes  
neurais artificiais**

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

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

Profa. Ana Rosa Fonseca de Aguiar Martins

Departamento de Engenharia Química e Materiais – PUC-Rio

Prof. André Luís Alberton  
Instituto de Química – UERJ

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

Rio de Janeiro, 26 de maio de 2022

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

### **Fernanda da Silva Pimentel**

Graduou-se em Engenharia Química na Universidade Federal Fluminense (UFF) em 2017. Atua profissionalmente na área de Propriedade Industrial.

#### Ficha Catalográfica

Pimentel, Fernanda da Silva

Desenvolvimento de modelos preditivos para a gaseificação de biomassa usando redes neurais artificiais / Fernanda da Silva Pimentel ; orientador: Florian Alain Yannick Pradelle ; coorientador: Brunno Ferreira dos Santos. – 2022.

134 f. : il. color. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Mecânica, 2022.

Inclui bibliografia

1. Engenharia Mecânica – Teses. 2. Otimização. 3. Gás de síntese. 4. Banco de dados. 5. Inteligência artificial. 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

Em primeiro lugar, agradeço à Deus pelo dom da vida.

Ao meu orientador, Florian, por estar sempre disposto a auxiliar e ensinar, com compreensão e paciência. Florian não mede esforços para o desenvolvimento de seus alunos.

Ao meu coorientador, Brunno, por todo o suporte e ensinamentos compartilhados, sempre com paciência e boa vontade.

À PUC-Rio, especialmente ao Departamento de Engenharia Mecânica, seu corpo docente e demais colaboradores.

Aos amigos e colegas do Programa de Pós-Graduação, pela companhia e pelas experiências trocadas, especialmente Isabela Fernandes Soares e João Gonçalves Neto, os quais foram fundamentais nessa caminhada.

Aos meus pais, Carlos Eduardo e Rosa Maria, e ao meu irmão, Guilherme, que sempre me apoiam e se doam ao nosso núcleo familiar com muito amor e carinho.

Ao meu esposo, Marcus Vinícius, que sempre incentiva meus projetos, com empatia, companheirismo e otimismo.

Ao Licks Attorneys por incentivar seus colaboradores à realização de pós-graduação.

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

## Resumo

Pimentel, Fernanda da Silva; Pradelle, Florian Alain Yannick (Orientador); dos Santos, Brunno Ferreira (Coorientador). **Desenvolvimento de modelos preditivos para a gaseificação de biomassa usando redes neurais artificiais**. Rio de Janeiro, 2022. 134p. Dissertação de Mestrado - Departamento de Engenharia Mecânica, Pontifícia Universidade Católica do Rio de Janeiro.

Na tentativa de reduzir os efeitos das emissões de dióxido de carbono, há uma necessidade por maior utilização de fontes de energia renováveis, tal como energia proveniente de biomassa. Para geração de energia a partir da biomassa, destaca-se o processo de gaseificação, por meio do qual é possível gerar um combustível nobre. Objetivou-se simular no software Matlab® a gaseificação da biomassa usando técnicas de inteligência artificial que são as Redes Neurais Artificiais (RNA). Particularmente, objetivou-se desenvolver modelos abrangentes de RNA com dez variáveis de entrada (carbono, hidrogênio, oxigênio, nitrogênio, material volátil, teor de umidade, cinzas, razão de equivalência, temperatura e razão vapor/biomassa), aplicáveis a uma diversidade de biomassa, com diversos tipos e concentrações de agentes de gaseificação em diferentes tipos de gaseificadores, capazes de prever a composição do gás de síntese (CO<sub>2</sub>, CO, CH<sub>4</sub> e H<sub>2</sub>). Para treinamento, teste e validação dos modelos, foram preparados bancos de dados robustos, a partir de informações coletadas em estudos anteriores disponíveis na literatura e do tratamento dos dados obtidos dos artigos. Foram avaliadas 33 topologias das redes neurais para eleger a melhor delas de acordo com quatro critérios referente a robustez do treinamento e do teste. A rede considerada como tendo a melhor topologia possui 10 neurônios na camada de entrada; 2 camadas intermediárias, com funções de ativação *logsig* e 10 neurônios em cada camada intermediária; função de ativação *purelin* na camada final; 4 neurônios na camada final; e algoritmo de treinamento *trainbr*. Tal rede possui um bom desempenho, com valores de R<sup>2</sup> de treinamento e de teste maiores que 0,88 e 0,70, respectivamente, para cada uma das quatro saídas. Para avaliação do modelo, uma validação foi executada, cujo desempenho não foi muito adequado, mas foi possível

identificar com uma métrica quantitativa simples as regiões mais confiáveis onde há uma maior densidade de dados no treinamento.

## **Palavras-chave**

Otimização, gás de síntese, banco de dados, inteligência artificial.

## Abstract

Pimentel, Fernanda da Silva; Pradelle, Florian Alain Yannick (Advisor); dos Santos, Brunno Ferreira (Co-advisor). **Development of predictive models for biomass gasification using artificial neural networks**. Rio de Janeiro, 2022. 134p. Dissertação de Mestrado - Departamento de Engenharia Mecânica, Pontifícia Universidade Católica do Rio de Janeiro.

In an attempt to reduce the effects of carbon dioxide emissions, there is a need for greater use of renewable energy sources, such as energy from biomass. In order to generate energy from biomass, the gasification process, by means of which it is possible to generate a noble fuel, can be highlighted. This work aimed to simulate the biomass gasification using artificial intelligence techniques, namely Artificial Neural Networks (ANN), using Matlab® software. Particularly, the objective was the development of ANN models with ten inputs (carbon, hydrogen, oxygen, nitrogen, volatile matter, moisture content, ash, equivalence ratio, temperature and steam/biomass ratio), applicable to a broad variety of biomass, with different types and concentrations of gasification agents in different types of gasifiers, capable of predicting the syngas composition (CO<sub>2</sub>, CO, CH<sub>4</sub> and H<sub>2</sub>). Robust databases were built for training, testing and validation of the models, based on information collected in previous studies available in the literature and on the treatment of data obtained from the papers. Thirty-three neural network topologies were evaluated in order to choose the best one according to four criteria regarding training and test robustness. The network considered to have the best topology has 10 neurons in the input layer; 2 hidden layers, with logsig activation functions and 10 neurons in each hidden layer; the activation function is purelin in the output layer; 4 neurons in the output layer; and the training algorithm is trainbr. Such network has a good performance, with R<sup>2</sup> values greater than 0.88 and 0.70 for training and test, respectively, for each of the four outputs. To evaluate the model, a validation was carried out, whose performance was not very appropriate, but it was possible to identify through a simple quantitative metric the more reliable regions where there is a greater density of training data.

## **Keywords**

Optimization, syngas, database, artificial intelligence.

## Sumário

1. Introdução .....	16
1.1. Contexto.....	16
1.2. Objetivos .....	19
1.3. Estrutura do presente trabalho.....	20
2. Revisão da Literatura .....	21
2.1. Processo de Gaseificação da Biomassa.....	21
2.2. Principais Parâmetros de Operação do Processo de Gaseificação da Biomassa .....	23
2.3. Tipos de Gaseificadores de Biomassa.....	26
2.3.1. Gaseificadores de Leito Fixo .....	27
2.3.2. Gaseificadores de Leito Fluidizado .....	31
2.3.3. Gaseificador de Leito Arrastado.....	34
2.4. Modelos para Simulação da Gaseificação da Biomassa .....	35
2.4.1. Modelos do tipo caixa transparente .....	35
2.4.2. Redes Neurais Artificiais (RNA).....	37
3. Metodologia.....	43
3.1. Bancos de Dados.....	43
3.2. Redes Neurais Artificiais (RNA) .....	48
4. Resultados e Discussão .....	53
4.1. Divisão dos dados de treinamento e de teste .....	53
4.2. Escolha da melhor topologia da RNA .....	55
4.3. RNA eleita e superfícies de resposta.....	61
4.3.1. Detalhamentos sobre a RNA eleita.....	61
4.3.2. Superfícies de resposta da RNA eleita .....	66
4.4. Validação do modelo.....	69
4.4.1. Análises qualitativa e quantitativa dos dados e perspectivas para melhoria dos modelos.....	72
4.4.1.1. Análise qualitativa.....	72
4.4.1.2. Análise quantitativa.....	78
4.5. Avaliação da robustez na região de validação.....	82
5. Conclusão .....	88
6. Referências Bibliográficas.....	90
Anexo I .....	97

Apêndice I - Trabalhos que utilizam RNA para simulação de gaseificação de biomassa .....	98
Apêndice II – Cálculo do LHV.....	102
Apêndice III – Cálculo do N <sub>2</sub> e demais gases.....	103
Apêndice IV – Histogramas das variáveis de entrada dos dados para treinamento, teste e validação das RNA .....	106
Apêndice V – Códigos Matlab utilizados no presente trabalho.....	116
Apêndice VI – Topologias das 33 redes neurais artificiais testadas.....	128
Apêndice VII – Análise quantitativa das 33 redes neurais artificiais testadas.....	129
Apêndice VIII - Valores dos pesos e bias da rede eleita .....	133

## Lista de Figuras

Figura 1 - Participação das fontes de energia renováveis e não renováveis na Oferta Interna de Energia (OIE) (fonte: EPE, 2021)	17
Figura 2 - Repartição da Oferta Interna de Energia (OIE) 2020 (fonte: EPE, 2021)	17
Figura 3 - Representação esquemática de um gaseificador contracorrente (fonte: adaptado de Sansaniwal, 2017)	27
Figura 4 - Representação esquemática de um gaseificador co-corrente (fonte: adaptado de Sansaniwal, 2017)	29
Figura 5 - Representação esquemática de um gaseificador de fluxo cruzado (fonte: adaptado de Sansaniwal, 2017)	30
Figura 6 - Representação esquemática de um gaseificador de leito fluidizado (fonte: livro Sánchez, 2010)	31
Figura 7 - Representação esquemática de um gaseificador de leito fluidizado borbulhante (fonte: Castro, 2020)	33
Figura 8 - Representação esquemática de um gaseificador de leito fluidizado circulante (fonte: Castro, 2020)	34
Figura 9 - Rede neural biológica (fonte: Aggarwal, 2018 – tradução livre)	37
Figura 10 - Rede neural artificial (fonte: Aggarwal, 2018 – tradução livre)	38
Figura 11 - (1) Perceptron com bias e (2) Rede de multicamadas com bias (fonte: Aggarwal, 2018 – tradução livre)	40
Figura 12 - Critérios para eleição da melhor topologia	50
Figura 13 - Histogramas dos dados de teste e treino para o H <sub>2</sub>	53
Figura 14 - Histogramas dos dados de teste e treino para o CO	53
Figura 15 - Histogramas dos dados de teste e treino para o CO <sub>2</sub>	54
Figura 16 - Histogramas dos dados de teste e treino para o CH <sub>4</sub>	54
Figura 17 - Histogramas representando as médias ponderadas entres os R <sup>2</sup> de teste e treino dos gases de saída para os 33 testes realizados	56
Figura 18 - Estrutura do modelo da RNA para predizer o gás de síntese a partir da gaseificação da biomassa	61
Figura 19 - Desempenho da melhor RNA para os dados de treinamento e teste – saída 1 (H <sub>2</sub> )	62
Figura 20 - Desempenho da melhor RNA para os dados de treino e teste – saída 2 (CO)	63
Figura 21 - Desempenho da melhor RNA para os dados de treino e teste – saída 3 (CO <sub>2</sub> )	64
Figura 22 - Desempenho da melhor RNA para os dados de treino e teste – saída 4 (CH <sub>4</sub> )	65
Figura 23 - Histogramas dos dados de treinamento do teor de umidade e temperatura	67
Figura 24 - Superfícies de resposta para as saídas: 1 - H <sub>2</sub> , 2 - CO, 3 - CO <sub>2</sub> e 4 - CH <sub>4</sub>	67

Figura 25 - Desempenho da melhor RNA para os dados de validação – saída 1 (H <sub>2</sub> )	70
Figura 26 - Desempenho da melhor RNA para os dados de validação – saída 2 (CO)	70
Figura 27 - Desempenho da melhor RNA para os dados de validação – saída 3 (CO <sub>2</sub> )	70
Figura 28 - Desempenho da melhor RNA para os dados de validação – saída 4 (CH <sub>4</sub> )	71
Figura 29 - Histogramas dos dados de treino e validação para o H <sub>2</sub>	72
Figura 30 - Histogramas dos dados de treino e validação para o CO	72
Figura 31 - Histogramas dos dados de treino e validação para o CO <sub>2</sub>	73
Figura 32 - Histogramas dos dados de treino e validação para o CH <sub>4</sub>	73
Figura 33 - Histogramas dos dados de treino e validação para o Carbono	74
Figura 34 - Histogramas dos dados de treino e validação para o Hidrogênio	74
Figura 35 - Histogramas dos dados de treino e validação para o Oxigênio	75
Figura 36 - Histogramas dos dados de treino e validação para o Nitrogênio	75
Figura 37 - Histogramas dos dados de treino e validação para o material volátil (VM)	75
Figura 38 - Histogramas dos dados de treino e validação para o teor de umidade (MC)	76
Figura 39 - Histogramas dos dados de treino e validação para cinzas	76
Figura 40 - Histogramas dos dados de treino e validação para ER	77
Figura 41 - Histogramas dos dados de treino e validação para temperatura	77
Figura 42 - Histogramas dos dados de treino e validação para S/B	77
Figura 43 - Gráficos distância x desvio das 1 <sup>a</sup> e 12 <sup>a</sup> menores distâncias	80
Figura 44 - Gráficos distância x desvio das 1 <sup>a</sup> a 12 <sup>a</sup> , 15 <sup>a</sup> , 20 <sup>a</sup> e 50 <sup>a</sup> menores distâncias	81
Figura 45 - Gráfico distâncias 7	83
Figura 46 - Subgrupo de validação	84
Figura 47 - Distância 7	85
Figura 48 - Superfícies de resposta para as quatro saídas, com base no subgrupo de validação próximo à região bem treinada	86

## Lista de Tabelas

Tabela 1 - Faixas de valores apontadas por Vassilev (2010) e dados excluídos dos bancos de dados para cada variável de entrada	47
Tabela 2 – $R^2$ para as quatro saídas	58
Tabela 3 - Topologias de 18 redes	59
Tabela 4 - Topologias das 6 redes finalistas	59
Tabela 5 – Cálculo do $N_p$ das 6 redes finalistas	60
Tabela 6 - Valores de $R^2$ e SSE de cada uma das saídas da validação	71
Tabela 7 – Ilustração da matriz distância de treinamento e validação	79
Tabela 8 – Descrição de diversos algoritmos de treinamento (adaptado de Aggarwal, 2005)	97

## Nomenclatura

b	Bias
BD	Banco de dados
BFB	Leito fluidizado borbulhante
CCE	Eficiência de conversão do carbono
CE	Camada de entrada
CF	Camada final
CFB	Leito fluidizado circulante
CFD	Fluidodinâmica computacional
CGE	Eficiência do gás frio
CH <sub>4</sub>	Metano
CI	Camada intermediária
CO	Monóxido de carbono
CO <sub>2</sub>	Dióxido de carbono
db	Base seca
ER	Razão de equivalência
FC	Carbono fixo
GY	Rendimento do gás
LHV	Poder calorífico inferior
logsig	Função sigmóide logística
MC	Teor de umidade
N <sub>p</sub>	Número de parâmetros total
N <sub>nCE</sub>	Número de neurônios da camada de entrada
N <sub>nCI</sub>	Número de neurônios da camada intermediária
N <sub>nCF</sub>	Número de neurônios da camada final
N <sub>b</sub>	Número de bias
purelin	Função de transferência linear
R <sup>2</sup>	Coeficiente de determinação
RNA	Redes Neurais Artificiais
S/B	Razão vapor/biomassa
SSE	Soma dos quadrados dos resíduos

T	Temperatura
tansig	Função tangente hiperbólica
VM	Material volátil
w	Peso

# 1. Introdução

## 1.1. Contexto

O total de emissões exercidas pelo homem associadas à matriz energética brasileira atingiu 398,3 milhões de toneladas de dióxido de carbono equivalente (Mt CO<sub>2</sub>-eq) em 2020, sendo a maior parte (179,8 Mt CO<sub>2</sub>-eq) gerada no setor de transportes (EPE, 2021). Na tentativa de reduzir os efeitos das emissões de dióxido de carbono (CO<sub>2</sub>), há uma necessidade por maior utilização de fontes de energia renováveis, tais como energia hidrelétrica, energia solar, energia eólica, energia proveniente de biomassa, energia oceânica e energia geotérmica. Particularmente com relação à biomassa, esta engloba toda a matéria orgânica vegetal, bem como resíduos orgânicos derivados de plantas, humanos, animais e vida aquática ou marinha. A biomassa é considerada uma fonte de energia renovável porque as plantas e colheitas crescem em escalas de tempo curtas em comparação com as escalas de tempo humanas (Michaelides, 2012). Nesse sentido, ressalta-se que o CO<sub>2</sub> emitido pela combustão de biomassa e seus derivados se enquadra em um ciclo em que há a sua reabsorção quando do crescimento de nova biomassa. Por outro lado, o CO<sub>2</sub> emitido pela combustão dos combustíveis fósseis não é absorvido de forma cíclica, o que aumenta sua concentração na atmosfera (Sánchez, 2010).

A matriz energética brasileira encontra-se em um patamar renovável superior ao observado no resto do mundo. Como pode ser observado na figura abaixo, as fontes de energia renováveis no Brasil representavam 48,4% e 46,1% da matriz energética em 2020 e 2019, respectivamente, enquanto o mundo registrava, em 2018, a participação de apenas 13,8% de energias renováveis (EPE, 2021).

### Participação das renováveis na OIE

Fonte: Agência Internacional de Energia (AIE) e EPE para o Brasil. Elaboração: EPE

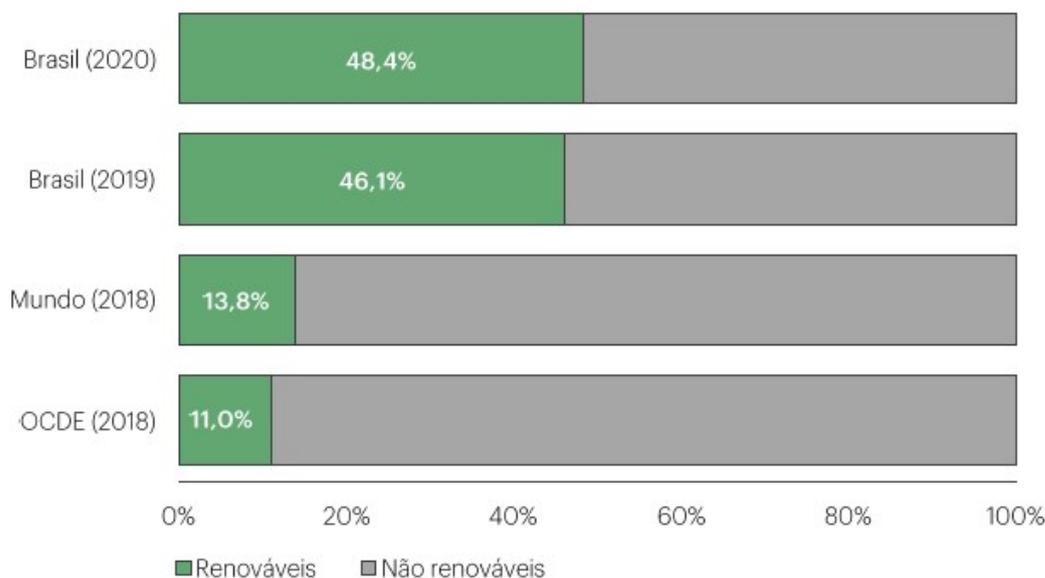


Figura 1 - Participação das fontes de energia renováveis e não renováveis na Oferta Interna de Energia (OIE) (fonte: EPE, 2021)

Além disso, dos 48,4% das fontes renováveis registrados em 2020, 19,1% equivalem à biomassa da cana-de-açúcar, 12,6% à hidráulica, 8,9% à lenha e carvão vegetal e 7,7% a outras fontes renováveis, sendo estas repartidas em lixívia, biodiesel, eólica, outras biomassas (incluindo casca de arroz, capim-elefante e óleos vegetais), solar, biogás e gás industrial de cravão vegetal, conforme ilustrado na Figura 2 abaixo (EPE, 2021).



Figura 2 - Repartição da Oferta Interna de Energia (OIE) 2020 (fonte: EPE, 2021)

O crescimento da contribuição das energias renováveis tem como força motriz o avanço da oferta de biomassa da cana-de-açúcar e biodiesel. Cabe ressaltar que o bagaço da cana-de-açúcar possui grande importância potencial como biocombustível. Dessa forma, pode-se concluir que a biomassa ocupa um papel importante no cenário energético brasileiro atual.

Com relação à geração de energia, é sabido que a pirólise, a gaseificação e a liquefação são abordagens comumente empregadas em processamento termoquímico a partir de biomassa para geração de biocombustíveis. Dentre essas tecnologias termoquímicas, a gaseificação da biomassa é considerada como sendo a mais eficiente e econômica (Sikarwar et al., 2017). Em linhas gerais, a gaseificação da biomassa pode ser definida como um processo complexo de conversão de uma matéria-prima sólida ou líquida em combustível gasoso útil e conveniente ou matéria-prima química que pode ser queimada para liberar energia ou usada para a produção de produtos químicos de valor agregado (Baruah, 2014). A gaseificação da biomassa é realizada na presença de um agente de gaseificação para produzir o gás de síntese (*syngas*), o qual contém monóxido de carbono, hidrogênio, dióxido de carbono e metano, com traços de alguns outros elementos. Assim, através do processo de gaseificação, é possível gerar um combustível nobre a partir da biomassa, de forma que este processo é considerado uma alternativa para as necessidades energéticas (Sánchez, 2010). O *syngas* atua também como matéria-prima para sínteses de biocombustíveis e produtos químicos e para geração de energia. Como exemplos da utilização do *syngas* pode-se listar a síntese de amônia, a produção de bio-hidrogênio e a geração de bio-metanol (Sikarwar et al., 2017). Com relação ao bio-metanol, de acordo com Eichler (2015), as principais reações de equilíbrio de síntese do referido combustível são as seguintes, sendo que a segunda ocorre em menor escala, em condições típicas de processo:



Para representar o processo de gaseificação da biomassa podem ser utilizados modelos matemáticos, os quais permitem prever o desempenho de um gaseificador para uma dada matéria-prima. A modelagem matemática atua como uma ferramenta importante para estudar o comportamento do gaseificador a fim de otimizar seu projeto e operação em comparação com a experimentação física e, assim, reduzir custo e tempo (Baruah, 2014). Dentre as diversas abordagens para

modelagem matemática relacionada ao processo de gaseificação, destaca-se a Rede Neural Artificial (RNA), a qual é uma técnica computacional biologicamente inspirada que imita o comportamento e o processo de aprendizagem do cérebro humano (Pandey, 2016). Uma RNA calcula uma função dos parâmetros de entrada, propagando os valores calculados dos neurônios de entrada para o(s) neurônio(s) de saída e usando os pesos como parâmetros intermediários (Aggarwal, 2018). Assim, como os estímulos externos são necessários para o aprendizado em organismos biológicos, o estímulo externo em RNA é fornecido pelos dados de treinamento que contêm exemplos de pares de entrada-saída da função a ser aprendida (Aggarwal, 2018).

As Redes Neurais Artificiais (RNA) são modelos de aprendizado baseados em redes neurais biológicas e podem ser supervisionados, uma vez que recebem um conjunto de dados de entrada e saída, e têm flexibilidade suficiente para modelar o mapeamento de entrada-saída não linear. Não há limitação ao tipo de conjunto de dados ou ao número de variáveis de entrada-saída. (Pandey, 2016). No entanto, é preciso ser muito diligente na aplicação da abordagem de RNA, já que se deve garantir que haja dados suficientes para formular o modelo. Uma rede neural pode retornar resultados ruins para dados que diferem dos dados originais com os quais foi treinada (Baruah, 2014).

A técnica de RNA tem sido usada com razoável sucesso para prever o rendimento e a composição do gás dos processos de gaseificação (Baruah, 2014). Nesse sentido, como será discutido em detalhes a seguir, no presente trabalho foram desenvolvidos modelos de RNA para prever as variáveis no estado final do processo de gaseificação da biomassa com boa precisão. Os modelos aqui desenvolvidos são abrangentes, posto que os dados utilizados nos mesmos englobam uma diversidade de tipo de biomassa, de condições experimentais (tipo e razão de agentes de gaseificação, temperatura) e de tipos de gaseificadores.

## 1.2. Objetivos

O principal objetivo do presente trabalho é desenvolver modelos de Redes Neurais Artificiais (RNA) robustos que sejam capazes de prever a composição do gás produzido no processo de gaseificação de uma ampla variedade de biomassa, com boa precisão. Objetiva-se modelos abrangentes, aplicáveis a uma diversidade

de biomassa, com diversos tipos e concentrações de agentes de gaseificação em diferentes tipos de gaseificadores.

Para a execução do presente trabalho, pode-se listar como objetivos secundários:

- coleta de informações, a partir de estudos anteriores disponíveis na literatura, para composição de bancos de dados;
- tratamento dos dados obtidos dos artigos para elaboração dos bancos de dados robustos a serem utilizados no treinamento e teste dos modelos, bem como para validação dos mesmos;
- validação do modelo de RNA eleito como tendo a melhor topologia;
- avaliação da robustez do modelo RNA através de uma validação.

### **1.3. Estrutura do presente trabalho**

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

Na seção 2, será apresentada uma revisão da literatura, com o intuito de aprofundar os conceitos relacionados ao processo de gaseificação de biomassa, incluindo principais parâmetros de operação e tipos de gaseificadores de biomassa, e modelos para simulação da gaseificação da biomassa, dando um destaque para 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 e de teste, a escolha da melhor topologia da RNA, a RNA eleita e superfícies de resposta, a validação do modelo – incluindo análises qualitativa e quantitativa dos dados e perspectivas para melhoria dos modelos – e, por fim, uma avaliação da robustez na região de validação.

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

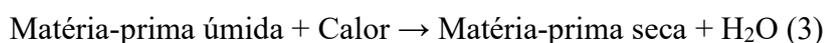
## 2. Revisão da Literatura

### 2.1. Processo de Gaseificação da Biomassa

A tecnologia de gaseificação começou a ser implementada há cerca de 100 anos e existem alguns dados de referência ao passar desses anos que merecem ser destacados. Na década de 1920, os carros na Suécia eram movidos a gaseificadores de madeira devido à ampla disponibilidade de biomassa de madeira e à falta de recursos petrolíferos. Durante a Segunda Guerra Mundial, entre 1939 e 1945, vários estudos foram realizados para otimizar os projetos de gaseificadores de madeira e melhorar seus desempenhos. Já nas décadas de 1970 e 1980, cerca de 40 empresas em todo o mundo se propuseram a construir plantas de gaseificação baseadas em biomassa, para gerar calor e energia (Sikarwar et al., 2017).

A gaseificação da biomassa pode ser definida como um processo complexo de conversão termoquímica altamente eficiente que converte diferentes matérias-primas de biomassa ou carvão mineral em um gás bruto composto principalmente por H<sub>2</sub>, CO, CO<sub>2</sub>, CH<sub>4</sub> e hidrocarbonetos leves que podem ser usados em outras aplicações, como combustível ou para a produção de produtos químicos (Sikarwar et al., 2017; Serrano, 2020). Por meio do referido processo, transformam-se diferentes tipos de resíduos em produtos valiosos para produção de energia e, dessa forma, a energia é produzida de forma renovável, além de reduzir o problema de descarte de resíduos. O processo de gaseificação é composto pelos seguintes processos térmicos: **(i)** secagem, **(ii)** pirólise, **(iii)** oxidação (ou combustão) e **(iv)** redução (Serrano, 2020).

Na etapa **(i)** de secagem, o teor de umidade da biomassa é reduzido. A referida etapa ocorre a uma temperatura de cerca de 100 a 200 °C (Puig-Arnavat, 2010). Devido à baixa temperatura, a biomassa não sofre decomposição térmica de suas matérias voláteis durante o processo de secagem, como indicado abaixo (Sansaniwal, 2017):



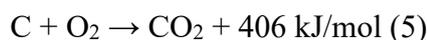
A pirólise (etapa **(ii)**) pode ser definida como a decomposição térmica da biomassa na ausência de oxigênio ou ar, por meio da qual o material volátil da

biomassa é reduzido. Dessa forma, ocorre a liberação de hidrocarbonetos em fase gasosa a partir da biomassa – os quais podem condensar a uma temperatura suficientemente baixa para gerar alcatrões líquidos – e carvão sólido. (Puig-Arnavat, 2010). De acordo com Sansaniwal (2017), a etapa de pirólise ocorre a uma temperatura de cerca de 125 a 500 °C e pode ser representada conforme segue:

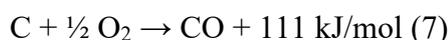


De acordo com Sánchez (2020), a pirólise ocorre entre 280 e 450 °C. No referido livro, é definido que a secagem e a pirólise da biomassa, com liberação de água, dos voláteis e do alcatrão, ocorrem em três fases: (1) evaporação da água; (2) decomposição dos carboidratos (hemicelulose, celulose e lignina)<sup>1</sup>, ocorrendo entre 250 e 300° C; e (3) produção de alcatrão e ácidos leves, a 350 a 430 °C. É ainda definido que os produtos obtidos são (1) gases não condensáveis (CO, CO<sub>2</sub>, H<sub>2</sub>, CH<sub>4</sub>, O<sub>2</sub>); (2) carvão vegetal; e (3) condensado: alcatrão e ácidos.

Com relação à etapa *(iii)* de oxidação, Puig-Arnavat (2010) menciona que esta trata de uma reação entre a biomassa sólida carbonizada e o oxigênio do ar, resultando na formação de CO<sub>2</sub>, e que o hidrogênio gerado durante a pirólise na biomassa também é oxidado para gerar água. Uma grande quantidade de calor é liberada com a oxidação do carbono e do hidrogênio, como pode ser observado nas equações de oxidação completa abaixo (Sansaniwal, 2017):



Ademais, se o oxigênio estiver presente em quantidades subestequiométricas, pode ocorrer oxidação parcial do carbono, resultando na geração de monóxido de carbono, como pode ser visto na equação abaixo (Puig-Arnavat, 2010; Susastriawan, 2017):




---

<sup>1</sup> Celulose, hemicelulose e lignina e extrativos são os principais componentes da biomassa (Kirubakaran, 2009).

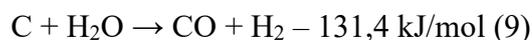
A temperatura de oxidação é de cerca de 800 a 1400 °C e o calor liberado durante a oxidação é usado para secagem, pirólise e outras reações endotérmicas durante a redução (Susastriawan, 2017). Ressalta-se que alguns dos principais parâmetros - como pressão e temperatura dentro do reator, tipo de agentes de gaseificação (oxigênio, ar e vapor) - desempenham um papel importante no rendimento do gás de síntese (Sansaniwal, 2017).

Por fim, o processo de redução (etapa **(iv)**) é onde ocorrem as principais reações de gaseificação. A faixa de temperatura do processo de redução é de cerca de 800 a 1000 °C. Os gases combustíveis no gás produzido são formados durante a redução por meio das seguintes reações endotérmicas (reação de Boudouard e reação de água-gás) e exotérmicas (reação de deslocamento água-gás e reação de formação do metano) (Susastriawan, 2017; Puig-Arnabat, 2010):

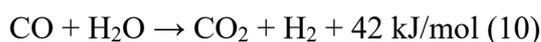
- Reação de Boudouard:



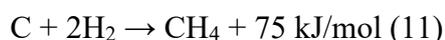
- Reação de água-gás:



- Reação de deslocamento água-gás:



- Reação de formação do metano:



## 2.2. Principais Parâmetros de Operação do Processo de Gaseificação da Biomassa

O produto final gasoso obtido pelo processo de gaseificação da biomassa é usualmente conhecido como gás de síntese ou *syngas*. O referido produto é composto principalmente por H<sub>2</sub>, CO, CO<sub>2</sub> e CH<sub>4</sub> (Sikarwar et al., 2017). Os principais fatores que afetam a composição dos gases de síntese e a eficiência da gaseificação são características da biomassa, tais como composição e teor de

umidade, e parâmetros do processo, tais como temperatura de gaseificação, agente de gaseificação, pressão e tipo de gaseificador. Tais fatores estão descritos com mais detalhes a seguir.

### **A) Composição da biomassa**

Com relação à composição, ressalta-se que toda biomassa tem, majoritariamente, carbono (C), hidrogênio (H) e oxigênio (O) e, por vezes, também nitrogênio (N) e enxofre (S). Tais elementos são usualmente reportados na literatura como última análise (*ultimate analysis*). Devido à sua composição, frequentemente a biomassa é representada pela fórmula geral  $C_xH_yO_z$ . Ressalta-se ainda que toda biomassa tem oxigênio (denominado bio-oxigênio) para converter, parcialmente ou totalmente, carbono em óxidos de carbono, *i.e.*, matéria combustível sólida em combustível gasoso. Todavia, o carbono também é consumido na formação do metano e do hidrocarboneto e na reação água-gás (Kirubakaran, 2009).

Não obstante, o teor de umidade (*moisture content* – MC), material volátil (*volatile matter* – VM), carbono fixo (*fixed carbon* – FC) e cinzas (*ash*), representando a fração inorgânica da biomassa, são comumente reportados na literatura como análise aproximada (*proximate analysis*) (Kirubakaran, 2009). Somente o material volátil e as moléculas de água associada a umidade do substrato contribuem a formação da mistura gasosa.

### **B) Teor de umidade da biomassa**

O teor de umidade da biomassa pode afetar a qualidade do gás produzido em termos do poder calorífico. Sabe-se que a biomassa com alto teor de umidade produz gás de síntese com baixo poder calorífico devido ao processo de pirólise incompleto durante a gaseificação. O calor necessário para essa etapa é insuficiente porque mais calor é absorvido pela matéria-prima para remover a umidade do substrato durante o processo de secagem (Susastriawan, 2017).

### C) Temperatura

No que tange a temperatura de gaseificação, destaca-se que, conforme esta aumenta, a porcentagem de gás combustível  $H_2$  e  $CO$  aumenta na composição da mistura e, portanto, o poder calorífico do gás produzido aumenta. Normalmente, o processo de oxidação tem a temperatura mais alta entre os outros processos, enquanto as temperaturas de todas os processos oscilam durante a gaseificação (Susastriawan, 2017).

### D) Agente de gaseificação

O agente de gaseificação pode ser ar, oxigênio ou vapor. O referido agente influencia a composição do *syngas* e, conseqüentemente, o poder calorífico. A gaseificação com ar gera um gás de síntese com um baixo poder calorífico devido ao alto teor de nitrogênio. A gaseificação com oxigênio, por sua vez, gera um gás de síntese com maior poder calorífico, quando comparado à gaseificação com ar, devido à ausência do nitrogênio. Ressalta-se, no entanto, que o uso de oxigênio aumenta drasticamente os custos operacionais e a complexidade da planta, devido à necessidade de uma unidade de separação de ar e dos sistemas de armazenamento e distribuição. Já a gaseificação a vapor produz um gás de síntese de alta qualidade (alto teor de hidrogênio e alto poder calorífico), mas precisa de calor, por isso é geralmente realizada com misturas de vapor/oxigênio (La Villetta, 2017). O vapor em forma combinada com ar resulta em maior concentração de hidrogênio, o que reduz a necessidade de energia de um sistema (Sansaniwal, 2017).

O ar é o mais utilizado como agente de gaseificação devido à sua disponibilidade e por consideração de custo (mais econômico). Na gaseificação com ar, um parâmetro de processo importante é a razão de equivalência (*equivalence ratio* - ER), a qual é definida como a razão entre a quantidade de ar usado na gaseificação e a quantidade de ar estequiométrico para a oxidação completa do carbono. Uma razão de equivalência típica para uma gaseificação efetiva está na faixa de 0,2 a 0,4. Valores menores que 0,2 significam que a gaseificação é dominada pela pirólise e valores maiores que 0,4 representam uma gaseificação onde predomina a combustão (Susastriawan, 2017).

Se o vapor for utilizado como o agente de gaseificação, uma quantidade adicional de hidrogênio pode ser adicionada ao produto de gaseificação a partir da molécula de H<sub>2</sub>O, devido ao fato da biomassa conter, em média, apenas 6% em peso de hidrogênio e por si só não ser uma fonte valiosa de hidrogênio. No entanto, o processo de craqueamento de H<sub>2</sub>O requer muita energia. Dessa forma, a gaseificação com vapor puro torna-se um processo altamente endotérmico, que exige fornecimento externo de energia. Em comparação, como anteriormente discutido, a gaseificação com ar produz uma mistura de gases com menor poder calorífico, mas não requer nenhuma entrada de energia externa, bem como nenhum custo adicional para geração de vapor. Todavia, para produzir uma mistura de gases com poder calorífico e teor de H<sub>2</sub> razoáveis, com um mínimo ou nenhuma entrada de energia externa, a gaseificação da biomassa pode ser realizada usando uma mistura de ar-vapor como agente de gaseificação. Nessa modalidade, as reações de oxidação exotérmica fornecem o calor para as reações de gaseificação endotérmicas (Loha, 2013).

No **Apêndice I** do presente trabalho, podem ser vistos exemplos de referências da literatura que utilizam ar, ar-vapor ou vapor como agente de gaseificação.

## **E) Pressão**

Com relação à pressão do gaseificador, tal reator pode trabalhar sob pressão atmosférica ou pressurizado (Puig-Arnavat, 2010). A quantidade de hidrocarbonetos aromáticos policíclicos (PAH) aumenta com o aumento da pressão do gaseificador, embora seja observada uma diminuição geral do alcatrão. Adicionalmente, sabe-se que pressões mais altas com razão de equivalência (*equivalence ratio* - ER) maiores resultam na diminuição da formação de hidrocarbonetos leves e de alcatrão (*tar*) devido a conversão completa de carbono (Sikarwar et al., 2017).

## **2.3. Tipos de Gaseificadores de Biomassa**

Em um processo de gaseificação de biomassa, o gaseificador deve ser capaz de produzir um gás combustível limpo e de alta qualidade. Por gás produzido

“limpo”, entende-se como tendo menor quantidade de impurezas, tais como alcatrão, material particulado,  $\text{SO}_x$ ,  $\text{NO}_x$  e  $\text{NH}_3$ . Quanto à qualidade do gás produzido, esta depende de vários fatores, como tipo de matéria-prima, configuração do reator, condições de operação e técnicas de limpeza e resfriamento do gás (Sansaniwal, 2017). Com relação à configuração, os gaseificadores mais comuns são classificados como **(i)** gaseificadores de leito fixo, **(ii)** gaseificadores de leito fluidizado e **(iii)** gaseificadores de leito arrastado.

### 2.3.1. Gaseificadores de Leito Fixo

Dentre os gaseificadores de leito fixo, existem o gaseificador contracorrente (*updraft*), o gaseificador co-corrente (*downdraft*) e o gaseificador de fluxo cruzado (*cross-draft*) (La Villetta, 2017), os quais serão detalhados a seguir. Como exemplos de trabalhos que utilizam o gaseificador de leito fixo, pode-se citar Plis (2011), Zhang (2015), Baruah (2017) e Trninić (2020).

#### A) Gaseificador Contracorrente (*updraft*)

O gaseificador contracorrente (*updraft*) é o mais simples dos gaseificadores e está ilustrado na Figura 3.

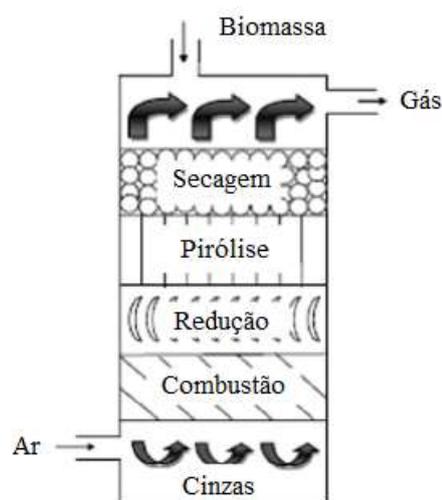


Figura 3 - Representação esquemática de um gaseificador contracorrente (fonte: adaptado de Sansaniwal, 2017)

O agente de gaseificação entra pela parte inferior do reator e move-se em direção à parte superior (fluxo ascendente), enquanto a biomassa entra em sentido contrário, *i.e.*, pelo topo do reator e move-se para baixo, pela ação da gravidade. No fundo do gaseificador, há uma grelha, onde a biomassa entra em combustão, o que gera um aumento da temperatura nesta zona inferior (zona de combustão ou oxidação). O gás quente, o qual tem baixo teor de oxigênio, sobe transportando calor para outras zonas do gaseificador, de modo que a biomassa que entrou pelo topo passa pela secagem, pirólise, redução e combustão (La Villetta, 2017).

Como vantagem do gaseificador contracorrente, pode-se mencionar sua eficiência térmica alta, consequente do fato dos gases quentes provenientes da zona de combustão pré-aquecerem a zona de combustível (biomassa) ao passarem por ela. Uma desvantagem do referido gaseificador é o fato de que produz um gás de síntese com muito vapor de água e alcatrão (Sánchez, 2010). No gaseificador contracorrente, a temperatura do gás de síntese na saída é baixa, em torno de 400 °C (Sánchez, 2010). O teor de alcatrão (*tar*), por sua vez, é alto no referido reator, uma vez que o gás não sofre reações de alta temperatura onde o alcatrão pode ser convertido através das reações de reforma a vapor do alcatrão (formando uma mistura de CO e H<sub>2</sub>O) e craqueamento do alcatrão (formando uma mistura de CO, CO<sub>2</sub>, CH<sub>4</sub> e H<sub>2</sub>) (La Villetta, 2017).

## **B) Gaseificador Co-corrente (*downdraft*)**

O gaseificador co-corrente (*downdraft*) pode ter uma configuração com topo aberto ou fechado. Nesse tipo de reator, para ambas as configurações, a biomassa e o agente de gaseificação têm o mesmo sentido de fluxo.

No gaseificador co-corrente de topo aberto, a biomassa e o agente de gaseificação são alimentados no topo do reator. O referido reator tem um topo aberto com ar forçado movendo-se para baixo por sucção para evitar formações de pontos quentes e, conseqüentemente, a fusão de cinzas. O fluxo de ar homogêneo reduz possíveis ineficiências no processo termoquímico. Esse tipo de reator possui alta flexibilidade e eficiência apreciável no processamento de materiais sólidos de natureza polidispersa, como casca de arroz de pequeno tamanho de partícula e baixa densidade (La Villetta, 2017).

No gaseificador co-corrente de topo fechado, a biomassa é alimentada pelo topo do reator e o agente de gaseificação é alimentado na zona de combustão, em fluxo descendente, passando pelas zonas de combustão e redução. A configuração desse gaseificador permite que os alcatrões e hidrocarbonetos voláteis oriundos da zona de pirólise passem pela zona de combustão, onde são craqueado pela alta temperatura, de modo que os alcatrões são convertidos em compostos mais leves (Sánchez, 2010). O gaseificador co-corrente de topo fechado pode ter duas configurações diferentes: um deles com reator cilíndrico reto e o outro com uma garganta no núcleo do reator, na zona de combustão, devido aos diferentes diâmetros das zonas de pirólise e de redução. Este último é conhecido como gaseificador Imbert (La Villetta, 2017) e está ilustrado na Figura 4.

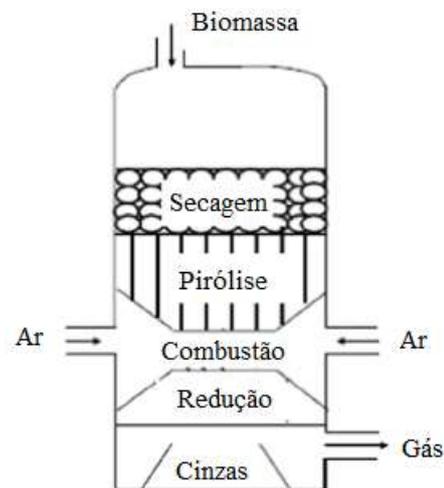


Figura 4 - Representação esquemática de um gaseificador co-corrente (fonte: adaptado de Sansaniwal, 2017)

Por último, ressalta-se que, como no gaseificador co-corrente o gás é liberado da zona de redução, este tende a ter quantidades de cinzas e fuligens significativas, diferentemente do gaseificador contracorrente em que tais partículas são filtradas quando o gás passa pelo combustível não queimado. Além disso, por não trocar calor diretamente com as zonas de pirólise e secagem, a temperatura do gás de síntese atinge valores em torno de 700 °C (Sánchez, 2010).

### C) Gaseificador de Fluxo Cruzado (*cross-draft*)

O gaseificador de fluxo cruzado (*cross-draft*) produz um gás combustível com propriedades intermediárias entre os gaseificadores contracorrente e co-corrente e está representado na Figura 5 abaixo.

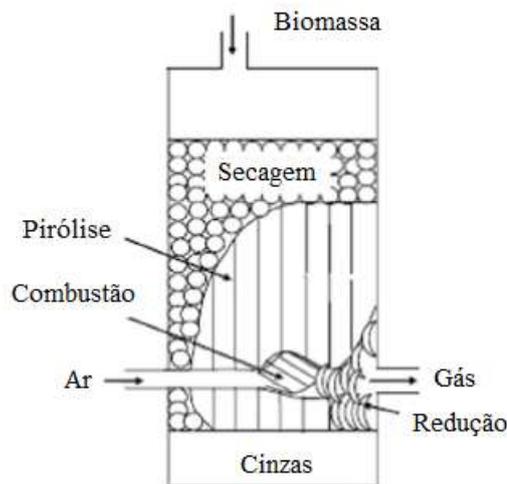


Figura 5 - Representação esquemática de um gaseificador de fluxo cruzado (fonte: adaptado de Sansaniwal, 2017)

Como pode ser observado na figura acima, o gaseificador de fluxo cruzado contém um injetor que direciona o agente de gaseificação no centro da zona de combustão. A zona de combustão e a zona de redução são concentradas no centro do gaseificador em um pequeno volume. A zona de combustão é caracterizada por temperaturas extremamente altas, já que a taxa de fornecimento de ar é consideravelmente maior que em outros tipos de gaseificadores. No gaseificador de fluxo cruzado, o gás de síntese é removido pela lateral, na mesma altura do leito em que é injetado o ar.

Como vantagens do gaseificador de fluxo cruzado, pode-se mencionar sua resposta rápida a variações de carga, seu peso reduzido e sua simplicidade de construção. No entanto, como desvantagem, tem-se que o referido reator é muito sensível a variações na composição e umidade do combustível (Sánchez, 2010).

### 2.3.2. Gaseificadores de Leito Fluidizado

Nos gaseificadores de leito fluidizado, o agente de gaseificação é soprado através de um leito de partículas sólidas a uma velocidade suficiente para manter as partículas em estado de suspensão (Puig-Arnavat, 2010). No interior do reator, é contida uma mistura de partículas de combustível e de material inerte, quando necessário. A referida mistura é suportada por uma placa distribuidora, através da qual o agente de gaseificação, previamente pressurizado, é obrigado a percorrer no sentido ascendente. É necessário que a vazão do referido agente de gaseificação seja tal que a força de arraste do escoamento do fluido supere o peso dos sólidos do leito e promova, paralelamente, a queima parcial do combustível (Sánchez, 2010). A Figura 6 a seguir ilustra um gaseificador de leito fluidizado.

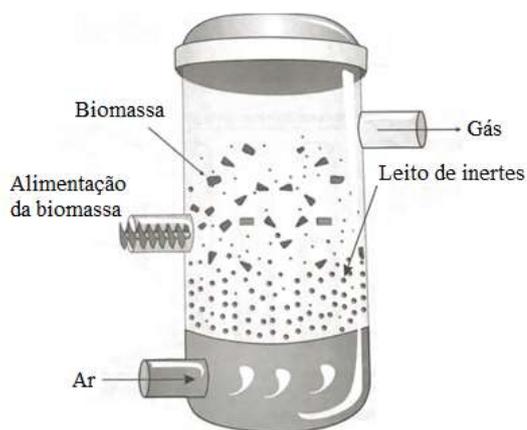


Figura 6 - Representação esquemática de um gaseificador de leito fluidizado (fonte: livro Sánchez, 2010)

Com relação ao material de leito inerte, utiliza-se mais comumente a sílica, podendo-se também utilizar outros sólidos a granel, como areia, olivina, contas de vidro, dolomita, etc. (Sansaniwal, 2017). As próprias cinzas do combustível sólido podem fazer o papel de material inerte em algumas situações, principalmente, quando utilizados carvões minerais com alto teor de impurezas minerais. Os objetivos de se adicionar material inerte são melhorar a qualidade da fluidização – por meio da intensificação do contato sólido-gás – e minimizar a formação de aglomerados na zona do leito, além do referido material atuar como um meio para intensificar a transferência de calor no escoamento gás-sólido (Sánchez, 2010).

Um leito fluidizado pode operar em condições quase isotérmicas e a temperatura de operação de um reator de leito fluidizado depende do ponto de fusão do material do leito, que geralmente varia de 800 a 900 °C. Tal faixa de temperatura é relativamente baixa e, portanto, as reações de gaseificação não atingem o equilíbrio químico em tais condições de baixa temperatura, a menos que algum catalisador seja usado. Outro fator responsável por não se atingir o equilíbrio químico é o curto tempo de residência do gás. Devido a esses fatores, os teores de hidrocarbonetos no gás de síntese no reator de leito fluidizado caem na faixa do gaseificador de leito fixo. Ressalta-se, no entanto, que a eficiência de conversão de carbono dos gaseificadores de leito fluidizado é comparativamente alta (até 95%) (Sansaniwal, 2017).

Dentre os gaseificadores de leito fluidizado, tem-se o gaseificador de leito fluidizado borbulhante (BFB) e o gaseificador de leito fluidizado circulante (CFB). Essa classificação está relacionada com as configurações das câmaras de reação dos gaseificadores. Em ambos os gaseificadores, as etapas de transformação da matéria prima a gás de síntese ocorrem praticamente de forma simultânea em toda a zona do leito, de modo que tais fases não são identificadas de modo separado no reator (Sánchez, 2010).

Os gaseificadores de leito fluidizado borbulhante (BFB) operam em regime fluidodinâmico borbulhante, em que é utilizada uma velocidade superficial do agente de gaseificação maior que a da mínima formação de bolha no leito. O referido leito é composto por duas fases: a emulsão e as bolhas, em que a primeira contém partículas sólidas com gás, enquanto as bolhas são praticamente livres de partículas sólidas (Sánchez, 2010). Nos gaseificadores de leito fluidizado borbulhante, a gaseificação de diferentes tipos de matéria-prima ocorre sob meio fluidizado de alta pressão. Geralmente, esses gaseificadores são projetados para operar a uma velocidade de gás muito baixa – normalmente abaixo de 1 m/s. As partículas sólidas enquanto se movem ao longo do fluxo de gás são separadas do gás no ciclone e são coletadas no fundo do reator de leito fluidizado. A maior parte do processo de conversão ocorre dentro da região do leito de borbulhamento (Sansaniwal, 2017). A Figura 7 a seguir ilustra um gaseificador de leito fluidizado borbulhante (BFB).

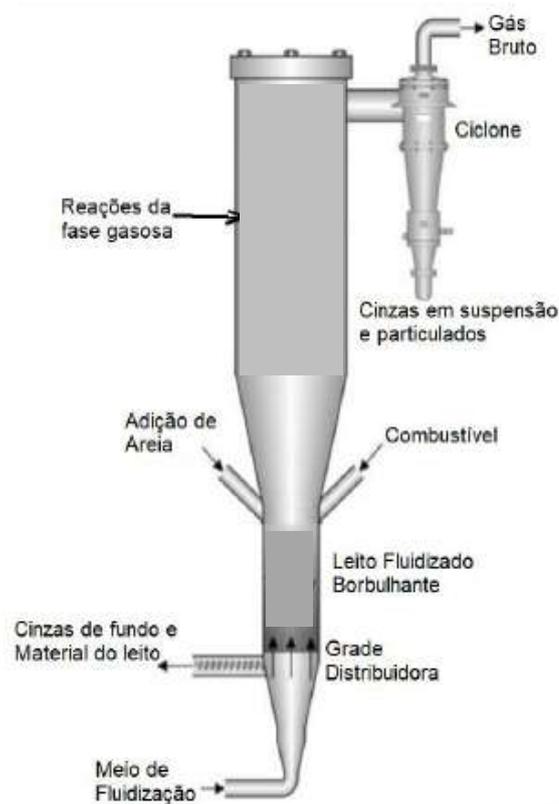


Figura 7 - Representação esquemática de um gaseificador de leito fluidizado borbulhante (fonte: Castro, 2020)

Nos gaseificadores de leito fluidizado circulante (CFB), o agente de gaseificação entra na base da coluna de fluidização a uma velocidade superficial acima da velocidade média de transporte das partículas e arrasta grande parte das partículas. O movimento de sólidos na direção axial e radial é intensificado com as altas velocidades do gás (variando de 3 a 10 m/s), gerando elevada transferência de calor e de massa na coluna de fluidização (Sansaniwal, 2017). O material sólido que é arrastado para fora da coluna de fluidização é capturado por um ciclone e, em seguida, despejado no fundo da coluna de retorno com escoamento de sólidos em sentido descendente. As partículas acumuladas no fundo da coluna de retorno são introduzidas novamente na coluna de fluidização, através do controle da válvula de recirculação. A recirculação dos sólidos tem o objetivo de aumentar o tempo de residência das partículas para que seja obtida uma máxima conversão do carbono contido na matriz sólida do combustível, resultando em ganho de rendimento e aumento da eficiência global do processo de gaseificação (Sánchez, 2010). A Figura 8 a seguir ilustra um gaseificador de leito fluidizado circulante (CFB).

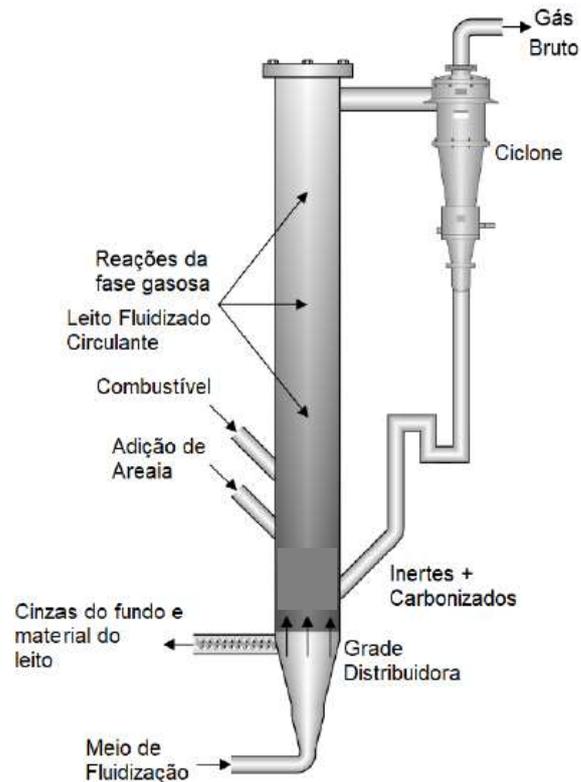


Figura 8 - Representação esquemática de um gaseificador de leito fluidizado circulante (fonte: Castro, 2020)

Como exemplos de trabalhos que utilizam o gaseificador de leito fluidizado, pode-se citar Sreejith (2013), Puig-Arnavat (2013), Pandey (2016), George (2018), Serrano (2020) e Christy (2021).

### 2.3.3. Gaseificador de Leito Arrastado

No gaseificador de leito arrastado (*entrained-flow*), partículas finas e oxidante (ar ou oxigênio) e/ou vapor são alimentados em escoamento de tipo co-corrente, o que resulta no oxidante e vapor rodeando ou arrastando as partículas sólidas conforme elas fluem através do gaseificador em uma nuvem densa. Tal gaseificador opera a altas condições de temperatura e pressão e é caracterizado por um fluxo extremamente turbulento que causa uma conversão rápida da matéria-prima e permite um alto rendimento (La Villetta, 2017). O gaseificador de leito arrastado é comumente utilizado para carvão, por poder ser alimentado com lama no modo de gaseificação direta, o que torna a alimentação de combustível sólido econômica em altas pressões (Puig-Arnavat, 2010). Todavia, esse tipo de gaseificador também

pode ser utilizado para biomassa, como pode ser visto, por exemplo, nos trabalhos de Zhou (2009) e Kibria (2019). O gaseificador de leito arrastado tem como características um curto tempo de residência, altas temperaturas, altas pressões e grandes capacidades (Puig-Arnavat, 2010).

## **2.4. Modelos para Simulação da Gaseificação da Biomassa**

A modelagem matemática tem um papel fundamental no estudo do comportamento da gaseificação da biomassa, com o objetivo de otimizar o referido processo. Por meio dos modelos matemáticos, é possível representar os fenômenos físicos e químicos que ocorrem no gaseificador, sendo considerados eficazes no fornecimento de orientação qualitativa sobre o efeito dos parâmetros de configuração, operação e matéria-prima no desempenho do gaseificador (Baruah, 2014).

Existem diversas abordagens para modelagem matemática do processo de gaseificação da biomassa. Dentre elas, são ressaltados os modelos de tipo caixa transparente, como os modelos considerando o equilíbrio termodinâmico, os modelos baseados na descrição da cinética e a fluidodinâmica computacional (CFD), bem como os modelos de tipo caixa preta, como as redes neurais artificiais (RNA).

### **2.4.1. Modelos do tipo caixa transparente**

Os modelos considerando o equilíbrio termodinâmico podem ser classificados como modelos estequiométricos – que se baseiam em constantes de equilíbrio – ou modelos não estequiométricos – que se baseiam no conceito de minimização da energia livre de Gibbs. No primeiro caso, as reações mais importantes são consideradas, enquanto outras reações são omitidas, motivo pelo qual podem haver erros na predição do modelo desenvolvido. No segundo caso, não é necessária a identificação e consideração das reações químicas, apenas o conhecimento da composição elementar, a temperatura e a pressão na gaseificação. Por meio dos modelos considerando o equilíbrio termodinâmico, é possível descrever o processo de gaseificação da biomassa com uma boa aproximação. Tal modelo é utilizado para prever a composição do gás produzido com base na

suposição de que os reagentes reagem em uma condição totalmente misturada por um período infinito de tempo. Ressalta-se, por fim, que os cálculos de equilíbrio termodinâmico são independentes da configuração do gaseificador, podendo ser mais adequados para estudos de processo sobre a influência dos parâmetros de processo e do combustível (biomassa) mais importantes. Apesar disso, os modelos considerando o equilíbrio têm sido utilizados com sucesso no estudo de diferentes tipos de gaseificadores (Baruah, 2014). Como exemplos de trabalhos que utilizam a metodologia de equilíbrio termodinâmico pode-se citar Habibollahzade (2021), Zainal (2001) e Jarunghammachote (2007).

Os modelos baseados na descrição da cinética consideram a modelagem matemática das equações que regem a cinética das reações de gaseificação dentro do gaseificador (a qual envolve o conhecimento da hidrodinâmica do leito e balanços de massa e energia para obter a composição em termo de gás, alcatrão e carvão em uma determinada condição operacional) e a hidrodinâmica do reator (a qual envolve o conhecimento do processo de mistura física). Tal modelo é considerado mais adequado e preciso em temperaturas de operação relativamente baixas do que o modelo de equilíbrio. Ademais, diferentemente dos modelos considerando o equilíbrio termodinâmico, os modelos baseados na descrição da cinética são utilizados para prever o rendimento do gás e a composição do gás produzido que um gaseificador alcança após um período finito de tempo. Essa modelagem é precisa e detalhada, mas é computacionalmente intensiva. Com o aumento da complexidade na configuração dos gaseificadores, a complexidade do modelo aumenta, tendo em vista que o equacionamento depende da hidrodinâmica do reator (Baruah, 2014). Como exemplos de trabalhos que utilizam modelo baseado na descrição da cinética pode-se citar Liu (2021), Wojnicka (2021) e Burra (2019).

A Fluidodinâmica Computacional (CFD) se baseia em soluções de um conjunto de equações simultâneas para conservação de massa, momento, energia e espécies em uma região discreta do gaseificador. O modelo de CFD é utilizado para prever a distribuição de temperatura, concentração e outros parâmetros dentro do reator. Tal modelo é considerado altamente preciso na previsão da temperatura e do rendimento de gás em torno do reator, se a fluidodinâmica do reator for bem conhecida. Por último, ressalta-se que a modelagem CFD da gaseificação de biomassa envolve a combinação de fluxo de particulado denso e química muito

específica, os quais são aspectos desafiadores na CFD (Baruah, 2014). Como exemplos de trabalhos que utilizam CFD pode-se citar Jahromi (2021) e Gao (2018). Os modelos de CFD incluem tipicamente modelos cinéticos.

#### 2.4.2. Redes Neurais Artificiais (RNA)

RNA pode ser definida como uma técnica computacional biologicamente inspirada que imita o comportamento e o processo de aprendizagem do cérebro humano (Pandey, 2016). O sistema nervoso humano contém neurônios (células), que são conectados uns aos outros com o uso de axônios e dendritos. As regiões de conexão entre axônios e dendritos são chamadas de sinapses. Em resposta aos estímulos externos, as forças das conexões sinápticas geralmente mudam, resultando no aprendizado dos organismos vivos (Aggarwal, 2018).

Uma Rede Neural Artificial calcula uma função dos parâmetros de entrada, propagando os valores calculados dos neurônios de entrada para o(s) neurônio(s) de saída e usando os pesos como parâmetros intermediários. Os pesos têm a função de conectar as unidades computacionais e atuam como os pontos fortes das conexões sinápticas em organismos biológicos. Cada entrada para um neurônio é dimensionada com um peso, que afeta a função calculada naquela unidade. (Aggarwal, 2018)

As figuras abaixo, extraídas do livro Aggarwal (2018) representam as conexões sinápticas entre os neurônios em uma rede neural biológica e uma rede neural artificial.

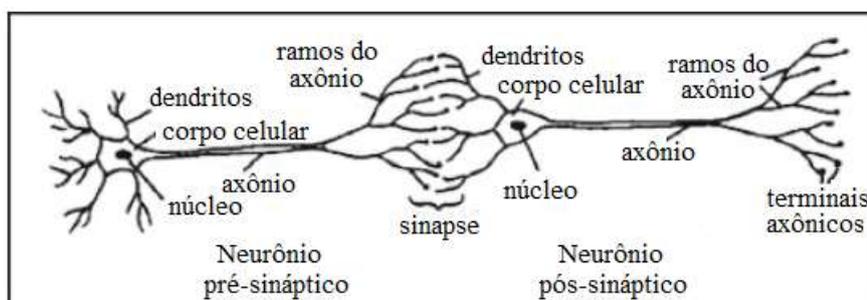


Figura 9 - Rede neural biológica (fonte: Aggarwal, 2018 – tradução livre)

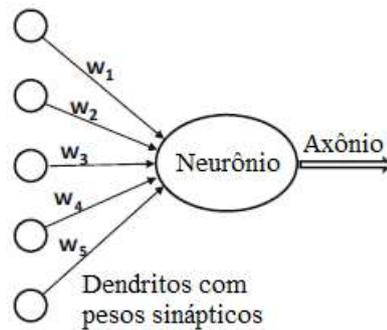


Figura 10 - Rede neural artificial (fonte: Aggarwal, 2018 – tradução livre)

Assim como os estímulos externos são necessários para o aprendizado em organismos biológicos, o estímulo externo em uma RNA é fornecido pelos dados de treinamento que contêm exemplos de pares de entrada-saída da função a ser aprendida (Aggarwal, 2018). Nesse sentido, uma RNA pode retornar resultados ruins para dados que diferem dos dados originais com os quais foi treinada, de forma que é necessário ser diligente na aplicação da abordagem de RNA (Baruah, 2014). O treinamento (ou aprendizado) pode ser definido como a tarefa de ajustar os parâmetros da RNA, conhecidos como pesos e bias. Existem diversos algoritmos de treinamento, que são utilizados para melhorar a capacidade de predição das variáveis de saída (Neto, 2021).

Nesse contexto, o artigo de Aggarwal, *et al* (2005) apresenta uma Tabela que descreve diferentes algoritmos de treinamento reproduzida no **Anexo I** do presente trabalho. O estudo descrito no referido artigo concluiu que os modelos de redes neurais treinadas usando o método de regularização Bayesiana (*trainbr*) forneceram os melhores resultados.

O algoritmo de treinamento *trainbr* atualiza os valores de peso e bias de acordo com a otimização de Levenberg-Marquardt e minimiza uma combinação de erros quadráticos e pesos para então determinar a combinação correta para produzir uma rede que generaliza adequadamente. O processo é chamado de regularização Bayesiana (*Bayesian Regularisation – BR*). Em outras palavras, o algoritmo *trainbr* pode ser definido como uma modificação do algoritmo de treinamento de Levenberg-Marquardt para produzir redes que generalizem bem, reduzindo a dificuldade de determinar a arquitetura de rede ideal. (Aggarwal, 2005; Kamble, 2015)

No método de regularização Bayesiana (trainbr), os pesos e os bias da rede são assumidos como variáveis aleatórias com distribuições especificadas. Os parâmetros de regularização estão relacionados às variâncias desconhecidas associadas a essas distribuições. O método de regularização Bayesiana envolve a modificação do índice de desempenho, o qual é dado pela seguinte equação (Kamble, 2015):

$$F = \beta \cdot E_d + \alpha \cdot E_w \quad (12)$$

em que:

$\alpha$  e  $\beta$ : parâmetros a serem otimizados (observação:  $\alpha$  e  $\beta$  não são independentes);

$E_d$ : soma média dos erros da rede ao quadrado;

$E_w$ : soma dos quadrados dos pesos da rede.

Uma característica do algoritmo de treinamento trainbr é que ele fornece uma medida de quantos parâmetros de rede (pesos e bias) estão sendo efetivamente utilizados pela rede (Kamble, 2015).

Com relação à arquitetura da RNA, há redes neurais de camada única (também chamada de *perceptron*) e redes neurais de multicamadas. Na rede de camada única, um conjunto de entradas é mapeado diretamente para uma saída usando uma variação generalizada de uma função linear. Já no caso de redes neurais multicamadas, os neurônios são organizados em camadas, em que as camadas de entrada e saída são separadas por um grupo de camadas ocultas ou intermediárias (*hidden layers*) (Aggarwal, 2018).

As arquiteturas do tipo *perceptron* de múltiplas camadas (MLP) constituem os modelos neurais artificiais mais utilizados e conhecidos, sendo construídas ordenando os neurônios em camadas e os sinais fluindo da camada de entrada para a camada de saída.

Com relação ao *perceptron*, considerando que:

(i) cada ponto de treinamento é da forma  $(X, y)$ , em que cada  $X = [x_1, \dots, x_d]$  contém  $d$  variáveis de característica e  $y \in \{-1, +1\}$  contém o valor observado (*i.e.*, valor fornecido) da variável de classe binária; e

(ii) a camada de entrada contém  $d$  nós que transmitem as  $d$  características  $X = [x_1 \dots x_d]$  com arestas de peso  $W = [w_1 \dots w_d]$  para um nó de saída; a saída predita pode ser definida pela seguinte equação (Aggarwal, 2018):

$$\hat{y} = \text{sign}\{\overline{W} \cdot \overline{X}\} = \text{sign}\left\{\sum_{j=1}^d w_j x_j\right\} \quad (13)$$

É válido ressaltar que a função linear é computada no nó de saída, de forma que o sinal deste valor real é utilizado para prever a variável dependente de  $X$ , como pode ser visto na equação (13) acima.

Não obstante, em algumas configurações, a abordagem acima não é suficiente para predição, por exemplo, quando a distribuição da classe binária é altamente desequilibrada. Nesses casos, é necessário incluir bias ( $b$ ), conforme segue (Aggarwal, 2018):

$$\hat{y} = \text{sign}\{\overline{W} \cdot \overline{X} + b\} = \text{sign}\left\{\sum_{j=1}^d w_j x_j + b\right\} \quad (14)$$

Dessa forma, bias pode ser definida como um parâmetro utilizado para ajustar a saída predita.

O bias pode ser incorporada como o peso de uma aresta usando um neurônio de bias. Isso é alcançado adicionando um neurônio que sempre transmite um valor de 1 para o nó de saída. O peso da aresta que conecta o neurônio de bias ao nó de saída fornece a variável bias. É importante mencionar que no caso de redes neurais de multicamadas, os bias podem ser usados nas camadas intermediárias e na camada de saída.

A Figura 11 ilustra (1) um *perceptron* com bias e (2) uma rede de multicamadas com bias. (Aggarwal, 2018)

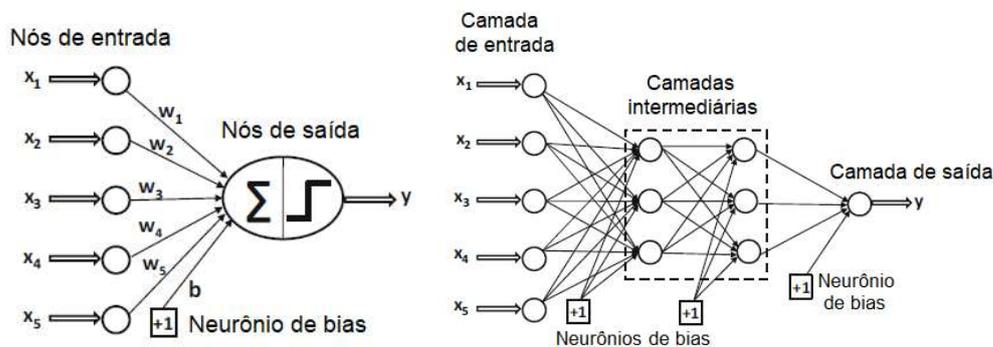


Figura 11 - (1) Perceptron com bias e (2) Rede de multicamadas com bias (fonte: Aggarwal, 2018 – tradução livre)

Com relação à função de sinal, esta atua como uma função de ativação (Aggarwal, 2018). A função de ativação pode ser definida como sendo a transformação não linear nos dados de entrada, permitindo a aprendizagem e execução de tarefas mais complexas. Enquanto os neurônios fazem uma transformação linear na entrada pelos pesos e bias, a função de ativação realiza a transformação não linear (Deep Learning Book, 2021).

Existem diversos tipos de função de ativação, dentre as quais são aqui destacadas a tangente hiperbólica (tansig), sigmóide logística (logsig) e função de transferência linear (purelin), indicadas abaixo (Neto, 2021):

- tansig:

$$f(x) = \frac{2}{(1 + e^{-2x})} - 1 \quad (15)$$

- logsig:

$$f(x) = \frac{1}{(1 + e^{-x})} \quad (16)$$

- purelin:

$$f(x) = x \quad (17)$$

Com relação à quantidade de neurônios a ser utilizada nas camadas intermediárias de uma rede neural artificial, é importante comentar os conceitos de *underfitting* e *overfitting*, que são problemas associados à escolha do número de neurônios. Se uma camada intermediária tiver poucos neurônios, pode haver um grande erro de treinamento devido ao *underfitting*. Dessa forma, o *underfitting* pode ser definido como um problema que ocorre quando há poucos neurônios que não conseguem detectar adequadamente os sinais em um conjunto complexo de dados. Se, por outro lado, uma camada intermediária tiver muitos neurônios, pode também haver um grande erro de treinamento devido ao *overfitting*. Desse modo, o *overfitting* pode ser definido como um problema que ocorre quando há muitos neurônios a serem treinados por um número limitado de informação contida no conjunto (Sheela, 2013).

Por último, alguns trabalhos que utilizam redes neurais artificiais para simulação de gaseificação de biomassa estão apresentados no **Apêndice I** do

presente trabalho. Para cada trabalho, são indicados o agente de gaseificação, o tipo de gaseificador, a biomassa utilizada, as variáveis de entrada e saída, e informações sobre RNA, tais como número de dados, divisão dos dados em percentual de treino, teste e validação, algoritmo de treinamento utilizado, função de ativação, número de camadas intermediárias e número de neurônios em cada camada. Como pode ser visto no **Apêndice I**, dentre os exemplos de trabalhos citados, predomina o algoritmo de treinamento *trainlm* e as funções de ativação utilizadas são *tansig*, *logsig* e *purelin*. Ressalta-se que os modelos dos referidos trabalhos são específicos para um agente de gaseificação, um tipo de gaseificador e/ou uma biomassa, de modo que nenhum deles simula um modelo aplicável a diversos tipos de agente de gaseificação, de gaseificador e de biomassa, como é feito no presente trabalho.

### 3. Metodologia

#### 3.1. Bancos de Dados

Foi realizada uma ampla busca por artigos científicos relacionados a gaseificação de biomassa, com o objetivo de coletar dados e desenvolver dois bancos de dados robustos e diversificados para serem utilizados na modelagem. O primeiro deles é o banco de dados principal, utilizado no treinamento e teste das RNA. O segundo se refere a um banco de dados para validação do modelo de RNA. Ambos os bancos de dados foram rigorosamente ajustados, como será detalhado a seguir.

Na montagem dos bancos de dados, foram coletadas as seguintes informações, quando aplicável:

- se os dados fornecidos pelo artigo são experimentais ou de simulação;
- tipos de gaseificador e de biomassa utilizados no estudo;
- agente de gaseificação;
- composição da biomassa (Carbono (C), Hidrogênio (H), Oxigênio (O), Nitrogênio (N) e/ou Enxofre (S));
- material volátil (VM);
- teor de umidade (MC);
- cinzas (Ash);
- carbono fixo (FC);
- razão de equivalência (ER);
- temperatura;
- material do leito;
- razão vapor/biomassa (S/B);
- gases de saída ( $H_2$ , CO,  $CO_2$ ,  $CH_4$  e/ou  $N_2$ );
- rendimento do gás (GY);
- eficiência do gás frio (CGE);
- eficiência de conversão do carbono (CCE); e
- LHV do gás.

Apesar de algumas dessas informações não terem sido utilizadas nos modelos desenvolvidos no presente trabalho, tais dados foram coletados para uso em trabalho posterior. Ressalta-se que alguns artigos forneciam dados numéricos por

meio de gráficos. Para leitura e extração dos dados numéricos dos referidos gráficos, utilizou-se a ferramenta online *WebPlotDigitizer* versão 4.5.

Não obstante, nem todos os documentos forneciam todas as informações desejadas. Por essa razão, fez-se necessário calcular algumas propriedades a partir das informações contidas nos artigos. A primeira delas foi o poder calorífico inferior (LHV) do gás. Para tal, utilizou-se os valores de referências encontrados em Green and Perry (2007). O cálculo realizado para o LHV está descrito no **Apêndice II** do presente trabalho. Outra propriedade calculada foi a concentração em base volumétrica de  $N_2$ , uma vez que alguns artigos forneciam os valores percentuais volumétricos de  $H_2$ ,  $CO$ ,  $CO_2$  e  $CH_4$  livres de  $N_2$ . Para esses dados, foram calculados os percentuais volumétricos de  $N_2$ , bem como dos outros gases considerando o  $N_2$ . O procedimento para realizar tal cálculo está disponível no **Apêndice III** do presente trabalho<sup>2</sup>.

Feito isso, foram criadas novas colunas no banco de dados com os valores percentuais dos gases corrigidos, considerando  $N_2$ , de forma que todos os gases estivessem na mesma base. Em seguida, foram calculados novos valores de LHV empregando os percentuais de  $H_2$ ,  $CO$  e  $CH_4$  que antes eram livres de  $N_2$ , mas agora consideram o mesmo. Tais valores corrigidos foram considerados na modelagem das RNA.

Com relação aos valores de razão de equivalência (ER) e de razão vapor/biomassa (S/B) que eventualmente não foram fornecidos por alguns artigos, as respectivas colunas foram preenchidas da seguinte maneira: ER preenchido como zero quando o agente de gaseificação é o vapor; e razão vapor/biomassa preenchida como zero quando o agente de gaseificação é o ar.

Não obstante, destaca-se que para casos em que as células do ER foram preenchidas com zero, pelo fato do agente de gaseificação ser o vapor, e o valor percentual em peso do N da biomassa ser considerado zero, o número de mols do  $N_2$  ( $n_{N_2}$ , conforme mencionado nos cálculos do **Apêndice III**) também é zero. A explicação para esse resultado é que, como o valor de N da biomassa é muito

---

<sup>2</sup> Nota: Um dos artigos utilizados apresentava dados em que os valores percentuais volumétricos de  $H_2$ ,  $CO$ ,  $CO_2$  e  $CH_4$  eram livres de nitrogênio, enquanto outros dados considerando o nitrogênio, sem uma indicação clara. Para o referido artigo, assumiu-se como critério: os casos em que as porcentagens do  $N_2$ , calculadas por diferença em relação à soma das porcentagens de  $H_2$ ,  $CO$ ,  $CO_2$  e  $CH_4$ , estavam abaixo de 25%, foram considerados como casos em que os valores de  $H_2$ ,  $CO$ ,  $CO_2$  e  $CH_4$  são livres de nitrogênio (assumindo um erro experimental refletido nos dados) e utilizou-se os cálculos indicados no **Apêndice III**.

pequeno (próximo ou igual a zero), poderia, por aproximação, considerar que todo nitrogênio que entra no reator viria do ar. Como nesses casos o agente de gaseificação é o vapor e não o ar, foi considerado que o %N<sub>2</sub> da saída do reator é igual a zero.

Em seguida, foi feita uma leitura cuidadosa do banco, a fim de eliminar dados incoerentes ou incompletos. Foram aplicados filtros nos dados, de modo a: **(i)** considerar apenas dados experimentais, desprezando-se os dados de simulação; **(ii)** considerar dados em que a coluna do VM estava preenchida e diferente de zero; **(iii)** excluir dados em que ER=S/B=0; **(iv)** excluir dados em que não havia informação sobre H<sub>2</sub>, CO, CO<sub>2</sub>, CH<sub>4</sub> e N<sub>2</sub>; **(v)** excluir dados em que não havia informação sobre temperatura; e **(vi)** excluir dados referentes a carvão mineral, uma vez que este está fora do escopo do presente trabalho<sup>3</sup>.

Além do acima exposto, outros ajustes foram necessários para padronização dos dados fornecidos pelos artigos, como, por exemplo, conversão de unidades e harmonização dos parâmetros da biomassa. Particularmente, a harmonização diz respeito a ajustes de base seca ou úmida e de base com ou sem cinzas.

Para a harmonização dos parâmetros da biomassa, foi feita uma verificação se há dados fora dos seguintes padrões: **(i)** soma percentual de C, H, O, N e S não sendo maior que 100%; e **(ii)** soma percentual de carbono fixo (FC), material volátil (VM), teor de umidade (MC) e cinzas não sendo maior que 100%. Como para alguns dos dados essas somas eram maiores que 100%, foi necessário revisar as informações trazidas pelos artigos utilizados nos bancos de dados.

Nesse sentido, notou-se que alguns artigos consideravam os elementos C, H, O, N e S com cinzas, enquanto outros não. Por isso, foi necessária uma homogeneização dos dados, em que quando os elementos C, H, O, N e S consideravam o teor de cinzas, ou seja, C + H + O + N + S + cinzas = 100%, foi feita a seguinte correção dos dados:

$$[\% \text{ C, H, O, N ou S}]_{\text{corrigido}} = [\% \text{ C, H, O, N ou S}]_{\text{artigo}} \times \frac{100}{100 - \text{cinzas}} \quad (18)$$

Além disso, notou-se que alguns artigos fornecem os valores de FC, VM e cinzas como sendo em base seca (*dry basis*), mas com valores de teor de umidade

<sup>3</sup> Com relação aos dados referentes a carvão mineral, foram excluídos dados em que o oxigênio da biomassa era igual a zero, o carbono da biomassa era igual a 99% e/ou o FC era igual a 84,40% (valores fornecidos por artigos da literatura).

(MC) diferentes de zero. Para esses casos, foi feito um ajuste dos dados da seguinte forma:

$$[\% \text{ FC, VM ou cinzas}]_{\text{corrigido}} = [\% \text{ FC, VM ou cinzas}]_{\text{artigo}} \times \frac{100 - \text{MC}}{100} \quad (19)$$

Desse modo, todos os parâmetros da biomassa foram padronizados.

Feito isso, os valores de cada parâmetro a ser utilizado como variável de entrada dos modelos de RNA – quais sejam: carbono (C), hidrogênio (H), oxigênio (O), nitrogênio (N), material volátil (VM), teor de umidade (MC), cinzas, razão de equivalência (ER), temperatura e razão vapor/biomassa (S/B) – foram cuidadosamente analisados e comparados com o que se espera a partir da literatura, de modo a excluir valores muito fora do comportamento usual.

Nesse contexto, o trabalho de Vassilev (2010), compila informações de valores dessas variáveis (exceto razão de equivalência (ER), temperatura e razão vapor/biomassa (S/B)) de diversas variedades de biomassa indicados por estudos anteriores, ou seja, o artigo Vassilev apresenta um conjunto de dados amplo da literatura, que pode ser tomado como base para uma comparação. No entanto, entende-se que os valores podem oscilar em torno das faixas indicadas por Vassilev, a se tratar de outras biomassas não consideradas pelo artigo.

A Tabela 1 abaixo sintetiza as faixas de valores apontadas pelo artigo Vassilev (2010) em amostras de diversos tipos de biomassa, bem como os dados excluídos dos bancos de dados do presente trabalho, tendo em vista os ensinamentos do referido artigo e análise dos dados. Com relação ao Nitrogênio, comenta-se que, apesar da referência Vassilev (2010) indicar uma faixa de valores para o N que vai até 12,2%, os dados compreendidos entre 6,13 e 6,81% do histograma gerado com os dados dos bancos foram desprezados, uma vez que havia um conjunto de quatro classes vazias entre 3,41 e 6,13%. Ademais, com relação ao teor de umidade, apesar da referência Vassilev (2010) indicar uma faixa de valores para o MC que vai até 62,9%, os dados acima de 18% foram desprezados, com o objetivo de diminuir a distorção dos dados observados no histograma gerado com os dados dos bancos. Por fim, destaca-se que, apesar das faixas apontadas por Vassilev (2010) para cinzas, tendo em vista a análise do histograma gerado com os dados dos bancos, os dados cujo valor de cinzas era de 36,53% foram desprezados para eliminar três classes vazias.

Tabela 1 - Faixas de valores apontadas por Vassilev (2010) e dados excluídos dos bancos de dados para cada variável de entrada

Variável de entrada	Faixa indicada (Vassilev, 2020)	Valores de dados excluídos dos bancos de dados
<b>Carbono (C), seco e livre de cinzas</b>	42,2 a 70,9%	77,55% e 79,34%
<b>Hidrogênio (H), seco e livre de cinzas</b>	3,2 a 11,2%	2,36%
<b>Oxigênio (O), seco e livre de cinzas</b>	16,4 a 49%	Nenhum
<b>Nitrogênio (N), seco e livre de cinzas</b>	0,1 a 12,2%	6,49 e 6,81%
<b>Material volátil (VM)</b>	30,4 a 79,7%, conforme medido em diferentes bases, ou de 47,8 a 86,3%, em base seca	7,77; 14,99; 20; 20,291; e 24,5
<b>Teor de umidade (MC)</b>	2,5 a 62,9%, sendo a média de 14,4%	0 e valores acima de 18% (particularmente, 19%; 21%; 22%; 23%; 23,5%; 25%)
<b>Cinzas</b>	0,1 a 43,3%, conforme medido em diferentes bases, ou de 0,1 a 46,3%, em base seca	36,53%
<b>Razão de equivalência (ER)</b>	---	Nenhum (tendo em vista apenas análise dos dados dos bancos de dados)
<b>Temperatura</b>	---	Valores entre 353 e 562 °C, e entre 911 e 1050 °C (tendo em vista apenas análise dos dados dos bancos de dados)
<b>Razão vapor/biomassa (S/B)</b>	---	4,04 (tendo em vista apenas análise dos dados dos bancos de dados)

Ao final, resultou-se em 603 dados totais, os quais estavam distribuídos em 491 dados no banco de dados principal, para treino e teste das RNA (cerca de 81,43 % dos dados totais) e 112 dados no banco de dados de validação (cerca de 18,57% dos dados totais). Como será discutido adiante, para desenvolvimento do modelo de RNA, os dados do banco principal foram divididos em 80% para treino (393 dados) e 20% para teste (98 dados). Portanto, considerando os dados totais, 65,17% foram destinados ao treinamento, 16,25% para teste e 18,58% para validação das RNA. Ressalta-se que o banco de dados de validação foi preparado de forma que os valores das variáveis de entrada e de saída estivessem compreendidos nas faixas de valores dos dados de treinamento. Para pronta visualização da distribuição dos dados, são aqui apresentados, no **Apêndice IV**, histogramas dos dados de treinamento, dados de teste e dados de validação das variáveis de entrada das RNA. Destaca-se que a distribuição dos histogramas em 8 grupos está definida para cada subconjunto, o que não garante que as classes tenham as mesmas fronteiras.

Os bancos de dados estão disponíveis sob solicitação aos autores do presente trabalho.

### 3.2. Redes Neurais Artificiais (RNA)

Os 491 dados do banco de dados principal foram separados em 80% para treino (393 dados) e 20% para teste (98 dados), utilizando o Matlab, versão R2020a (vide código denominado “Primeira parte” do **Apêndice V**).

Em seguida, os dados de treino e de teste foram alimentados no programa principal para geração dos modelos de Redes Neurais Artificiais (vide código denominado “RNA principal” no **Apêndice V**). Os modelos continham dez variáveis de entrada, quais sejam: carbono (C), hidrogênio (H), oxigênio (O), nitrogênio (N), material volátil (VM), teor de umidade (MC), cinzas, razão de equivalência (ER), temperatura e razão vapor/biomassa, e quatro saídas, quais sejam: CO<sub>2</sub>, CO, CH<sub>4</sub> e H<sub>2</sub> (composição do gás de síntese).

Ressalta-se que o mesmo grupo de treino, bem como o mesmo grupo de teste, foi utilizado para cada topologia de rede, de modo que todos os testes realizados do programa principal rodaram com o mesmo grupo de dados de treino e o mesmo de teste. É importante mencionar ainda que foi feita uma análise das divisões dos dados de treino e teste antes de proceder com os testes dos modelos de RNA. Para a

referida análise, foram gerados histogramas dos dados de treino e teste para as saídas H<sub>2</sub>, CO, CO<sub>2</sub> e CH<sub>4</sub>, com o objetivo de avaliar como estava a divisão dos dados de treinamento e teste, *i.e.*, se os mesmos estavam balanceados.

Foram realizados 33 testes, variando-se a topologia das redes, *i.e.*, o número de camadas intermediárias, o número de neurônios em cada uma das camadas, as funções de ativação das camadas intermediárias e da camada final e o algoritmo de treinamento. Foram testadas redes com uma a quatro camadas intermediárias, com número de neurônios variando de 10 a 400 nas referidas camadas. As funções de ativação testadas foram *logsig* e *tansig* nas camadas intermediárias e *purelin*, *tansig* e *satlin* na camada final. Os algoritmos de treinamento testados no presente trabalho foram: *trainbr*, *trainlm*, *traingdx*, *traingd*, *traingda*, *traingdm*, *trainbfg*, *traincgb*, *traingcf*, *traingcp* e *trainrp*. No entanto, a partir da leitura de trabalhos anteriores, era esperado que os algoritmos *trainbr* e *trainlm* trariam os melhores resultados. Foram feitas diversas combinações dos parâmetros a fim de buscar a melhor topologia da RNA. Destaca-se que como havia quatro saídas - H<sub>2</sub>, CO, CO<sub>2</sub> e CH<sub>4</sub> - as redes testadas compreendiam 4 neurônios na camada final.

Ao realizar-se cada um dos testes, foram registrados os valores de R<sup>2</sup> e SSE de treino e teste de cada saída (análise quantitativa do modelo).

Nesse sentido, a precisão do modelo pode ser estimada pelo coeficiente de determinação (R<sup>2</sup>), de acordo a seguinte equação (Neto, 2021):

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_{observado} - Y_{predito})^2}{\sum_{i=1}^n (Y_{observado} - Y_{m\u00e9dio})^2} \quad (20)$$

Por meio dos valores de R<sup>2</sup> (coeficiente de determinação), é possível avaliar se o modelo faz um bom ajuste dos dados. O R<sup>2</sup> varia de 0 a 1, de modo que quanto mais perto de 1 (ou 100%), melhor é o ajuste de dados pelo modelo.

O SSE (*Sum of Squared Errors* - Soma dos quadrados dos resíduos), por sua vez, é a soma das diferenças ao quadrado entre o valor observado e o valor predito, de tal maneira que quanto menor é o SSE, melhor é o resultado. O SSE pode ser calculado de acordo com a seguinte equação (Neto, 2021):

$$SSE = \sum_{i=1}^n (Y_{observado} - Y_{predito})^2 \quad (21)$$

Após registro dos valores de  $R^2$  e SSE de treino e teste de cada saída, tirou-se a média ponderada entre os valores de  $R^2$  de treino e de teste, bem como de SSE de treino e SSE de teste, para cada uma das saídas, de modo a obter uma métrica única para facilitar a visualização do desempenho das redes.

Foi feita uma análise de cada uma das topologias e a escolha das melhores delas teve como base quatro critérios assumidos. A Figura 12 a seguir ilustra a metodologia de escolha das topologias.

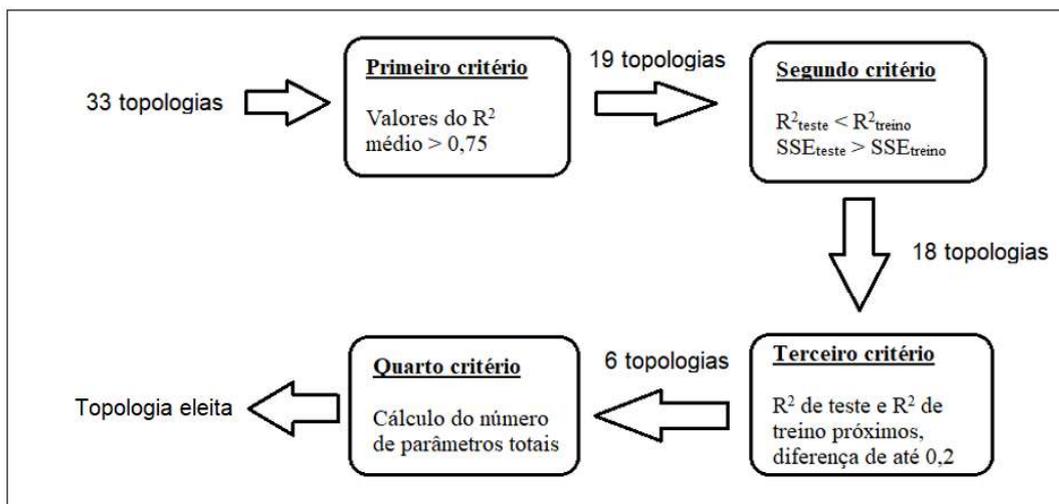


Figura 12 - Critérios para eleição da melhor topologia

O primeiro critério diz respeito à seleção das topologias em que os valores do  $R^2$  médio do CO (e de todos os gases de saída) eram maiores que 0,75. O segundo critério está relacionado à análise se os valores de  $R^2$  de teste são menores que  $R^2$  de treino, bem como se os valores de SSE de teste são maiores que SSE de treino. Cabe ressaltar que é aceitável que  $R^2_{teste}$  seja maior que  $R^2_{treino}$  e que  $SSE_{teste}$  seja menor  $SSE_{treino}$  quando a diferença entre eles é pequena. O terceiro critério se refere à análise se os valores de  $R^2$  de teste e  $R^2$  de treino são próximos, assumindo como aceitável uma diferença entre  $R^2$  de treino e de teste de até 0,2 (este valor foi estipulado após observação das diferenças entre  $R^2$  de teste e  $R^2$  de treino) para garantir uma uniformidade entre os resultados do treino e do teste. Por último, o quarto critério refere-se ao cálculo da quantidade de parâmetros das seis redes finalistas.

Com relação ao último critério, é importante ressaltar que quanto menor a quantidade de parâmetros, menor é o tempo que a rede leva para processar, sendo,

portanto, a melhor escolha. Para o cálculo do número de parâmetros total ( $N_p$ ), leva-se em conta o número de neurônios em cada camada e os bias associados. O  $N_p$  pode ser calculado da seguinte forma:

$$N_p = (N_{n_{CE}} \times N_{n_{CI}} + N_b) + (N_{n_{CI}} \times N_{n_{CI}} + N_b) + (N_{n_{CI}} \times N_{n_{CF}} + N_b) \quad (22)$$

em que:

$N_p$  é o número de parâmetros total;

$N_{n_{CE}}$  é o número de neurônios da camada de entrada;

$N_{n_{CI}}$  é o número de neurônios da camada intermediária;

$N_{n_{CF}}$  é o número de neurônios da camada final;

$N_b$  é o número de bias, em que cada bia acompanha um neurônio.

O número de parâmetros total da RNA não pode passar o número de dados de treino (393 dados). Quando isso ocorre, significa que há mais parâmetros a serem treinados do que o número de dados de treino e há chance de o modelo indicar uma predição errada, por exemplo, podendo ter uma situação de *overfitting*. Dessa forma, nessa situação em que o número de parâmetros total é maior do que o número de dados de treino, significa que há muito neurônios ou muitas camadas na rede, em outras palavras, há um dimensionamento superior ao ideal.

No entanto, quando se utiliza o algoritmo de treinamento *trainbr*, este tem parada antecipada. É possível confiar no modelo que utiliza o *trainbr* mesmo quando o número de parâmetros total é superior ao número de dados do treino. A confiança atrelada ao algoritmo de treinamento *trainbr* ocorre porque tal algoritmo pode apresentar um número de parâmetros efetivos menor do que o número de parâmetros totais. Ao usar o *trainbr*, é reduzida a possibilidade de *overfitting*.

Em um caso, por exemplo, onde compara-se duas redes com o mesmo número de camadas e neurônios em cada uma das camadas, diferindo pelo algoritmo de treinamento, em que a primeira rede utiliza *trainbr* e a segunda rede um outro algoritmo (por exemplo, *trainlm*), a rede que utiliza o *trainbr* tem a possibilidade de trabalhar com um número de parâmetros efetivos menor (critério de confiança), caracterizando essa rede como tendo uma melhor topologia. Se for utilizado outro algoritmo de treinamento, por exemplo, *trainlm*, não é possível confiar no modelo se o número de parâmetros total é maior que o número dados treino, uma vez que o referido modelo pode indicar uma predição errada.

Após eleição da melhor topologia de RNA, foi feita montagem e análise de gráficos de superfícies de resposta para cada uma das quatro saídas. Para geração dos referidos gráficos, foi utilizado o Matlab, versão R2022a (código denominado “Superfícies de resposta” no **Apêndice V**). As variáveis de entrada utilizadas nas superfícies de resposta foram teor de umidade e temperatura e foram utilizados os dados de treinamento da RNA.

Adicionalmente, foi realizada a validação do modelo de RNA desenvolvido, utilizando o Matlab R2022a (vide código denominado “Validação RNA” do **Apêndice V**) e o segundo banco de dados previamente preparado contendo 112 dados. Ressalta-se que os referidos dados não estão contidos no banco de dados principal, que foi utilizado para treino e teste.

## 4. Resultados e Discussão

### 4.1. Divisão dos dados de treinamento e de teste

De modo a analisar as divisões dos dados totais nos grupos de treino e teste, antes que os mesmos fossem utilizados nos testes dos modelos de RNA, foram gerados os histogramas reproduzidos a seguir (Figuras 13 a 16) dos grupos de dados de treino e teste para cada uma das quatro saídas, H<sub>2</sub>, CO, CO<sub>2</sub> e CH<sub>4</sub>.

A partir da Figura 13, nota-se que, para o treinamento, a região mais populosa do H<sub>2</sub> (*i.e.*, região de maior aprendizagem da rede) está compreendida ente 4,05 e 43,17% e este perfil também é observado para os dados de teste.

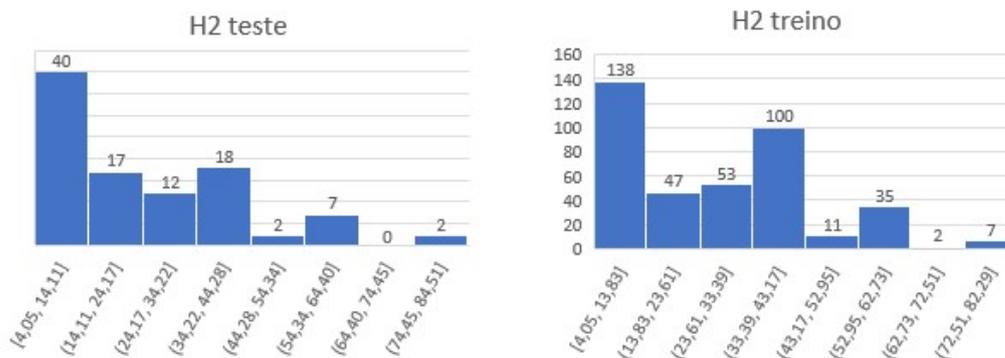


Figura 13 - Histogramas dos dados de teste e treino para o H<sub>2</sub>

Com relação ao CO, a partir da Figura 14, nota-se que, para o treinamento, a região mais populosa está compreendida ente 5,80 e 23,09%. A região mais populosa dos dados de teste do CO está englobada nessa faixa.

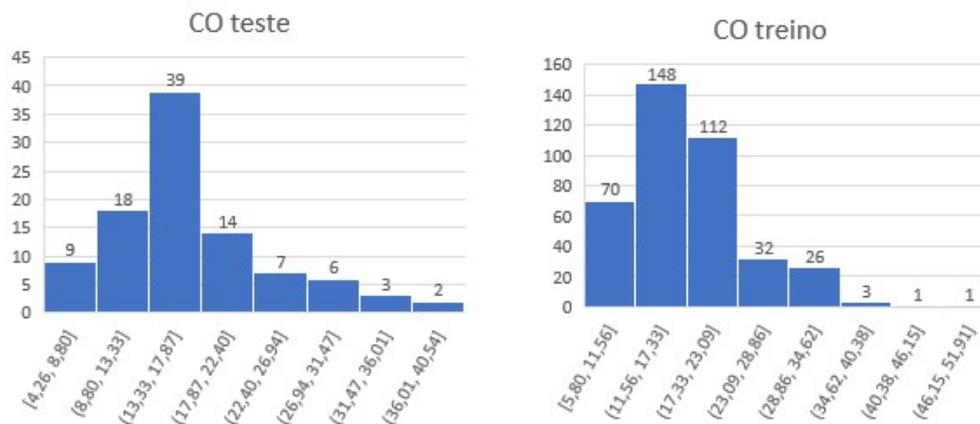


Figura 14 - Histogramas dos dados de teste e treino para o CO

Com relação ao CO<sub>2</sub>, a partir da Figura 15, nota-se que, para o treinamento, a região mais populosa está compreendida ente 11,22 e 29,92% e este perfil também é observado para os dados de teste.

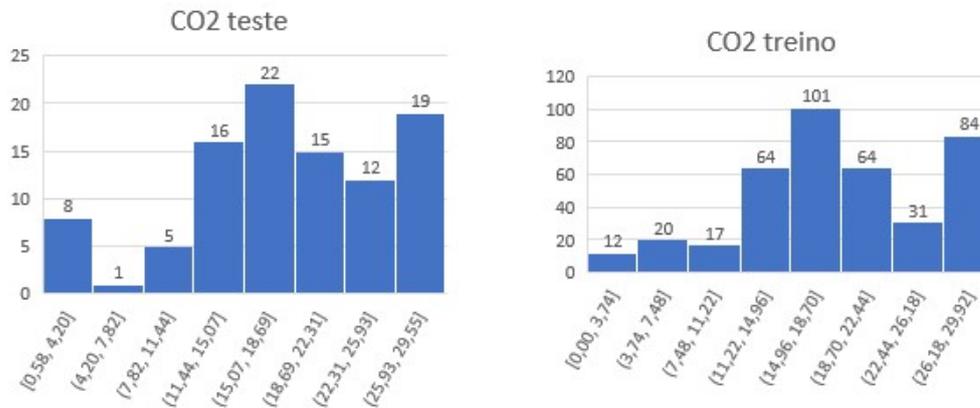


Figura 15 - Histogramas dos dados de teste e treino para o CO<sub>2</sub>

Por fim, com relação ao CH<sub>4</sub>, a partir da Figura 16 é possível observar que, para o treinamento, a região mais populosa está compreendida ente 0 e 12,01%. Esta faixa de valores, em que a rede seria mais bem treinada, engloba quase toda a faixa de dados do teste.

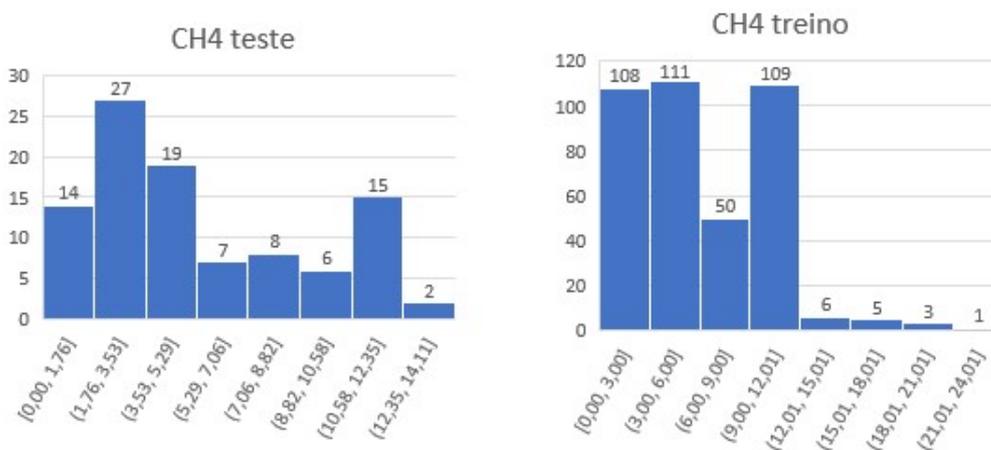


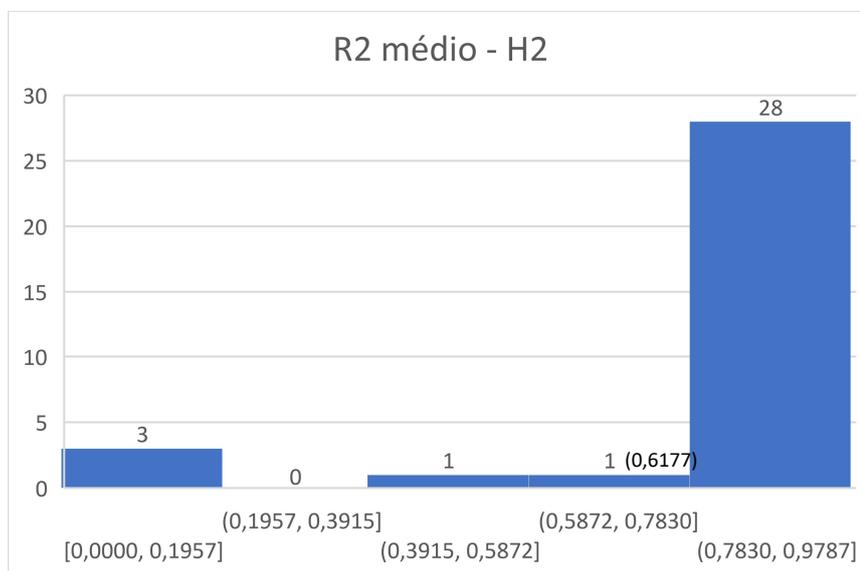
Figura 16 - Histogramas dos dados de teste e treino para o CH<sub>4</sub>

Portanto, como os dados de teste/treino estavam bem distribuídos, a divisão dos referidos dados foi mantida para utilização nos modelos de Redes Neurais Artificiais.

## 4.2. Escolha da melhor topologia da RNA

Foi feita uma análise das 33 topologias testadas (indicadas no **Apêndice VI**) para eleger a melhor, através dos critérios descritos na seção metodologia. A análise quantitativa das redes encontra-se no **Apêndice VII**. Vale comentar que algumas células referentes ao SSE não estão preenchidas, tendo em vista que se tratam de casos em que os valores de  $R^2$  são muito baixos e por isso é irrelevante considerar o SSE.

- **Primeiro critério:** Foram selecionadas as topologias em que os valores do  $R^2$  médio do CO eram maiores que 0,75. Ao analisar os valores do  $R^2$  médio para cada gás, notou-se que o CO apresentava maior quantidade  $R^2$  médio menor que 0,75. Essa conclusão pode ser ilustrada nos histogramas gerados das médias ponderadas entres os  $R^2$  de treino e teste de cada um dos gases de saída. Como pode ser observado nas figuras abaixo, 28 dados estavam acima 0,75 para o  $H_2$ , 19 dados estavam acima 0,75 para o CO, 26 dados estavam acima 0,75 para o  $CO_2$ , e 25 dados estavam acima 0,75 para o  $CH_4$ .



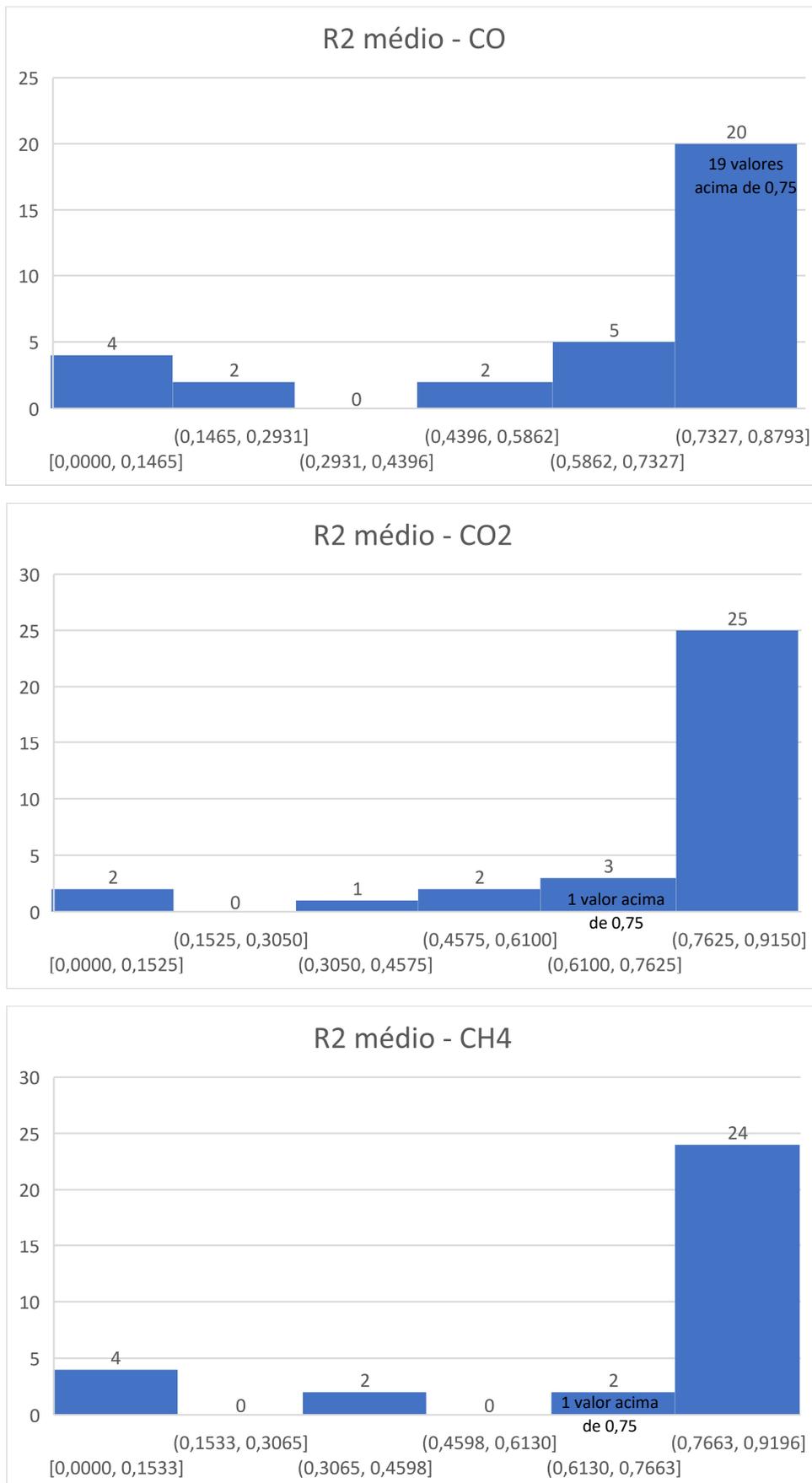


Figura 17 - Histogramas representando as médias ponderadas entres os R<sup>2</sup> de teste e treino dos gases de saída para os 33 testes realizados

Além disso, notou-se que para valores de  $R^2$  médio maiores que 0,75, todos os outros três gases também apresentavam  $R^2$  médio maior que 0,75. Com a seleção de topologias em que os valores do  $R^2$  médio do CO são maiores que 0,75, foram desprezados 14 testes (testes de números 1 a 3, 9 a 17, 26 e 27), resultando em 19 topologias.

É válido comentar que foi observado que as topologias cujos algoritmos de treinamento eram `traingdx`, `traingd`, `traingda`, `traingdm`, `trainbfg`, `traincgb`, `traingcf`, `traingcp` e `trainrp` foram descartadas nesse momento (testes 9 a 17). Além disso, as topologias cuja função de ativação eram `satlin` também foram descartadas nesse critério (testes 26 e 27). Os testes 1 a 3 também desprezados no presente critério possuíam o menor número de neurônio (10) e apenas uma camada intermediária.

- **Segundo critério:** Análise se os valores de  $R^2$  de teste são menores que  $R^2$  de treino, bem como se os valores de SSE de teste são maiores que SSE de treino.

**a)** Considera-se que  $R^2_{\text{teste}} < R^2_{\text{treino}}$ , no entanto é aceitável o inverso quando a diferença entre ambos é pequena, por exemplo, menor que 0,05. Em todas as 19 topologias resultantes do primeiro critério,  $H_2$ , CO e  $CO_2$  atendem ao critério de que  $R^2_{\text{teste}} < R^2_{\text{treino}}$ . Com relação ao  $CH_4$ , os testes 8, 21 e 31 não atendem a esse critério, no entanto, nos três casos, a diferença entre  $R^2_{\text{teste}}$  e  $R^2_{\text{treino}}$  é menor que 0,05. Dessa forma, nenhum teste foi desprezado.

**b)** Considera-se que  $SSE_{\text{teste}} > SSE_{\text{treino}}$ , no entanto é aceitável o inverso quando a diferença entre ambos é pequena, por exemplo, quando as SSE de treino e teste estão na mesma ordem de grandeza. No teste 25, a  $SSE_{\text{teste}}$  é significativamente menor do que a  $SSE_{\text{treino}}$  para as saídas  $H_2$  e  $CO_2$ , e ligeiramente menor para a saída  $CH_4$  (mas mantendo a mesma ordem de grandeza). Essa topologia foi eliminada. Ressalta-se que esta foi a única topologia em que a função de ativação utilizada na camada final foi `tansig`. Nos testes 5 e 20, a  $SSE_{\text{teste}}$  é menor do que a  $SSE_{\text{treino}}$  para a saída  $H_2$ , todavia, as referidas topologias não foram desprezadas. Resultou-se então em 18 topologias, após exclusão do teste 25.

- **Terceiro critério:** Análise se os valores de  $R^2$  de teste e  $R^2$  de treino são próximos. No critério anterior, considerou-se que  $R^2_{\text{teste}} < R^2_{\text{treino}}$ , uma vez que o inverso - *i.e.*,  $R^2_{\text{teste}} > R^2_{\text{treino}}$  - indica que o modelo está em uma faixa de valores (sendo aceitável uma diferença pequena, por exemplo, menor que 0,05). Além dessa consideração, é ainda mais coerente que os  $R^2$  de treino e de teste sejam mais próximos possíveis. Dessa forma, foi assumido como aceitável uma diferença entre  $R^2$  de treino e de teste de até 0,2. Na tabela 2 abaixo, estão indicadas as diferenças entre os  $R^2$  das 18 topologias resultantes do segundo critério, em que os valores destacados em vermelho são aqueles em que a diferença é maior que 0,2.

Tabela 2 –  $R^2$  para as quatro saídas

Teste	H2			CO			CO2			CH4		
	R <sup>2</sup> treino1	R <sup>2</sup> teste1	Diferença	R <sup>2</sup> treino2	R <sup>2</sup> teste2	Diferença	R <sup>2</sup> treino3	R <sup>2</sup> teste3	Diferença	R <sup>2</sup> treino4	R <sup>2</sup> teste4	Diferença
4	0,9794	0,8857	0,0937	0,8939	0,5246	0,3693	0,9259	0,8012	0,1247	0,9035	0,8856	0,0179
5	0,9747	0,8895	0,0852	0,8890	0,6125	0,2765	0,9243	0,7722	0,1521	0,9003	0,8520	0,0483
6	0,9755	0,9119	0,0636	0,8857	0,7003	0,1854	0,9274	0,8135	0,1139	0,9063	0,8581	0,0482
7	0,9760	0,8563	0,1197	0,8799	0,5791	0,3008	0,9250	0,8301	0,0949	0,9063	0,8861	0,0202
8	0,9838	0,9491	0,0347	0,9033	0,7564	0,1469	0,9307	0,8159	0,1148	0,9169	0,9248	-0,0079
18	0,9866	0,3855	0,6011	0,9307	0,2563	0,6744	0,9671	0,0866	0,8805	0,9372	0,2657	0,6715
19	0,9846	0,9343	0,0503	0,9086	0,6893	0,2193	0,9348	0,7904	0,1444	0,9203	0,9060	0,0143
20	0,9835	0,9286	0,0549	0,9003	0,7168	0,1835	0,9290	0,8225	0,1065	0,9168	0,8808	0,0360
21	0,9847	0,9508	0,0339	0,9058	0,7244	0,1814	0,9328	0,7980	0,1348	0,9184	0,9242	-0,0058
22	0,9805	0,9168	0,0637	0,8871	0,7150	0,1721	0,9269	0,7846	0,1423	0,9042	0,8770	0,0272
23	0,9839	0,8765	0,1074	0,9078	0,6375	0,2703	0,9347	0,7969	0,1378	0,9196	0,9016	0,0180
24	0,9847	0,9548	0,0299	0,9080	0,7640	0,1440	0,9354	0,8333	0,1021	0,9212	0,9024	0,0188
28	0,9835	0,9264	0,0571	0,9009	0,6162	0,2847	0,9311	0,8186	0,1125	0,9149	0,8713	0,0436
29	0,9822	0,9046	0,0776	0,9000	0,5815	0,3185	0,9292	0,8354	0,0938	0,9160	0,9014	0,0146
30	0,9841	0,9330	0,0511	0,9040	0,6364	0,2676	0,9315	0,8108	0,1207	0,9191	0,8950	0,0241
31	0,9826	0,9201	0,0625	0,8932	0,6565	0,2367	0,9289	0,8157	0,1132	0,9144	0,9301	-0,0157
32	0,9828	0,9087	0,0741	0,8993	0,6257	0,2736	0,9301	0,8038	0,1263	0,9146	0,8461	0,0685
33	0,9876	0,2155	0,7721	0,9362	0,0327	0,9035	0,9691	0,0098	0,9593	0,9389	0,0259	0,9130

Assim, os testes 4, 5, 7, 18, 19, 23 e 28 a 33 foram desprezados, resultando em seis topologias finalistas, quais sejam: 6, 8, 20, 21, 22 e 24. É importante ressaltar os testes 18 e 33 para os quais a diferença das quatro saídas foram maiores que 0,2 (particularmente, maiores que 0,6, ou seja, uma diferença consideravelmente alta, tendo em vista o critério assumido). Tais testes foram os únicos, dentre os 18 analisados no presente critério, cujo algoritmo de treinamento utilizado foi o trainlm, como pode ser visto na Tabela 3 abaixo. Os outros 16 testes foram feitos com o algoritmo de treinamento trainbr.

Tabela 3 - Topologias de 18 redes

Teste	Topologia da rede										
	Número de neurônios				Função de ativação					CF	Algoritmo de treinamento
	1ª CI	2ª CI	3ª CI	4ª CI	1ª CI	2ª CI	3ª CI	4ª CI			
4	20	N/A	N/A	N/A	tansig	N/A	N/A	N/A	purelin	trainbr	
5	10	10	N/A	N/A	logsig	tansig	N/A	N/A	purelin	trainbr	
6	10	10	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr	
7	10	10	N/A	N/A	tansig	tansig	N/A	N/A	purelin	trainbr	
8	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr	
18	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainlm	
19	40	40	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr	
20	60	60	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr	
21	40	40	40	N/A	logsig	logsig	tansig	N/A	purelin	trainbr	
22	30	30	30	30	logsig	logsig	tansig	tansig	purelin	trainbr	
23	20	60	N/A	N/A	tansig	logsig	N/A	N/A	purelin	trainbr	
24	60	80	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr	
28	100	N/A	N/A	N/A	tansig	N/A	N/A	N/A	purelin	trainbr	
29	100	N/A	N/A	N/A	logsig	N/A	N/A	N/A	purelin	trainbr	
30	80	80	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr	
31	200	N/A	N/A	N/A	tansig	N/A	N/A	N/A	purelin	trainbr	
32	400	N/A	N/A	N/A	tansig	N/A	N/A	N/A	purelin	trainbr	
33	400	N/A	N/A	N/A	tansig	N/A	N/A	N/A	purelin	trainlm	

Nota-se, por exemplo, que a diferença entre as topologias dos testes 8 e 18 é apenas o algoritmo de treinamento utilizado (respectivamente, trainbr e trainlm), uma vez que ambas possuem mesmo número de camadas intermediárias, mesmo número de neurônios em cada camada e mesmas funções de ativação. A rede 8 possui diferenças entre  $R^2$  de treino e teste menores que 0,2 para todos os gases de saída, enquanto na rede 18 essas diferenças são maiores que 0,6, como pode ser observado na Tabela 2.

- **Quarto critério:** consistiu no cálculo do número de parâmetros total ( $N_p$ ) das seis redes finalistas (testes 6, 8, 20, 21, 22 e 24), cujas topologias estão indicadas abaixo.

Tabela 4 - Topologias das 6 redes finalistas

Teste	Topologia da rede										
	Número de neurônios				Função de ativação					CF	Algoritmo de treinamento
	1ª CI	2ª CI	3ª CI	4ª CI	1ª CI	2ª CI	3ª CI	4ª CI			
6	10	10	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr	
8	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr	
20	60	60	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr	
21	40	40	40	N/A	logsig	logsig	tansig	N/A	purelin	trainbr	
22	30	30	30	30	logsig	logsig	tansig	tansig	purelin	trainbr	
24	60	80	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr	

Tendo em vista essas topologias, foram realizados os seguintes cálculos dos números de parâmetros totais:

Tabela 5 – Cálculo do Np das 6 redes finalistas

Rede	Cálculo do Np	Np
<b>6</b>	$(10 \times 10 + 10) + (10 \times 10 + 10) + (10 \times 4 + 4)$	<b>264</b>
<b>8</b>	$(10 \times 20 + 20) + (20 \times 20 + 20) + (20 \times 4 + 4)$	<b>724</b>
<b>20</b>	$(10 \times 60 + 60) + (60 \times 60 + 60) + (60 \times 4 + 4)$	<b>4.564</b>
<b>21</b>	$(10 \times 40 + 40) + (40 \times 40 + 40) + (40 \times 40 + 40) + (40 \times 4 + 4)$	<b>3.884</b>
<b>22</b>	$(10 \times 30 + 30) + (30 \times 4 + 4)$	<b>3.244</b>
<b>24</b>	$(10 \times 60 + 60) + (60 \times 80 + 80) + (80 \times 4 + 4)$	<b>5.864</b>

Com relação às redes 20, 21, 22 e 24, apesar de também utilizarem o algoritmo de treinamento trainbr, elas possuem valores de Np muito maiores do que o número de dados de treino (cerca de 10 vezes maior), de modo que devem ser desconsideradas. Como pode ser observado na tabela acima, as redes com os menores números de parâmetros totais são as redes 6 e 8, em que os valores de Np são, respectivamente, 264 e 724. Com relação à rede 8, apesar de seu Np ser maior que o número de dados de treino, a referida rede utiliza o algoritmo de treinamento trainbr, de modo que pode trabalhar com um número de parâmetros efetivos menor do que seu número de parâmetros total, conforme explicado anteriormente. Além disso, cabe ressaltar que o Np da rede 8 tem a mesma ordem de grandeza do número de dados de treino. Portanto, considera-se que a rede 8 apresenta uma boa topologia. Já a rede 6 é considerada como tendo a melhor topologia por possuir número de parâmetros total ( $N_p = 264$ ) menor que o número de dados de treino (393 dados).

A rede 6, eleita como tendo a melhor topologia dentre as 33 redes analisadas, possui duas camadas intermediárias, função de ativação logsig nas referidas camadas, 10 neurônios em cada uma das camadas intermediárias, função de ativação purelin na camada final, algoritmo de treinamento trainbr, 10 neurônios na camada de entrada e 4 neurônios na camada final. A Figura abaixo representa uma estrutura do modelo da melhor RNA.

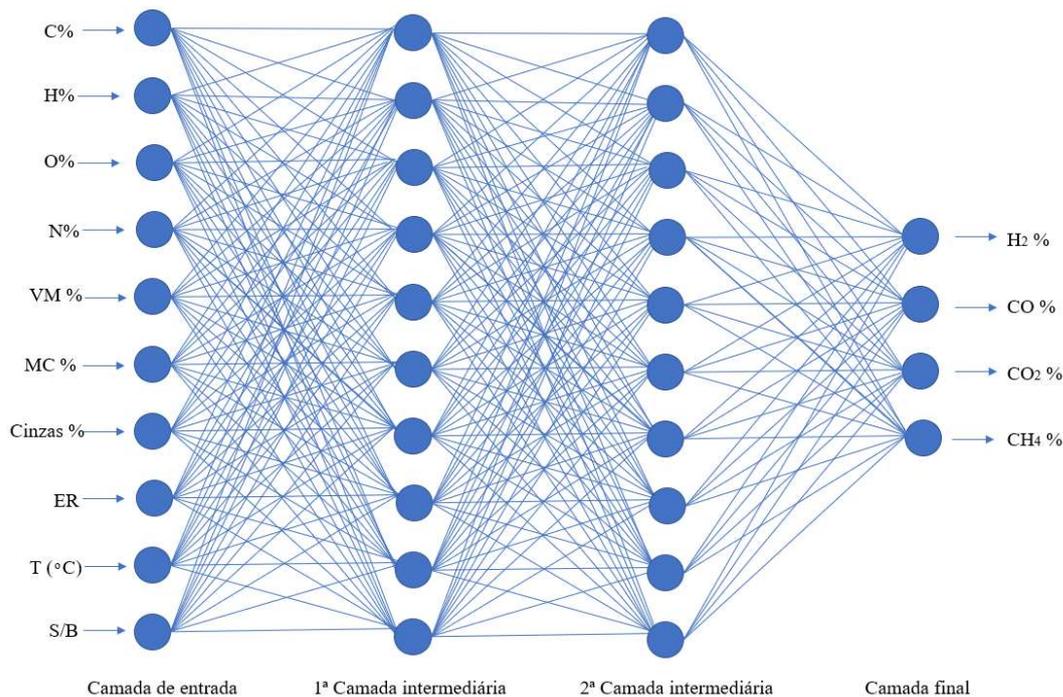


Figura 18 - Estrutura do modelo da RNA para prever o gás de síntese a partir da gaseificação da biomassa

Os valores dos pesos e bias da referida rede podem ser encontrados no **Apêndice VIII**.

### 4.3. RNA eleita e superfícies de resposta

#### 4.3.1. Detalhamentos sobre a RNA eleita

Além do acima exposto, são aqui apresentados os gráficos das saídas (“*outputs*”) versus o alvo (“*target*”), bem como de saídas real e predita para cada amostra de dado ao executar a rede 6 para os dados de treinamento e teste para as concentrações de H<sub>2</sub>, CO, CO<sub>2</sub> e CH<sub>4</sub> (Figuras 19 a 22 abaixo). Nos gráficos superiores das Figuras abaixo, o alvo ou *target T* representa as concentrações de H<sub>2</sub>, CO, CO<sub>2</sub> ou CH<sub>4</sub> do banco de dados, enquanto a saída (“*Output Y*”) é a predição feita pela RNA do alvo T. O ajuste linear é a regressão linear entre *Output Y* e *Target T* e sua forma ideal seria  $Y = T$ .

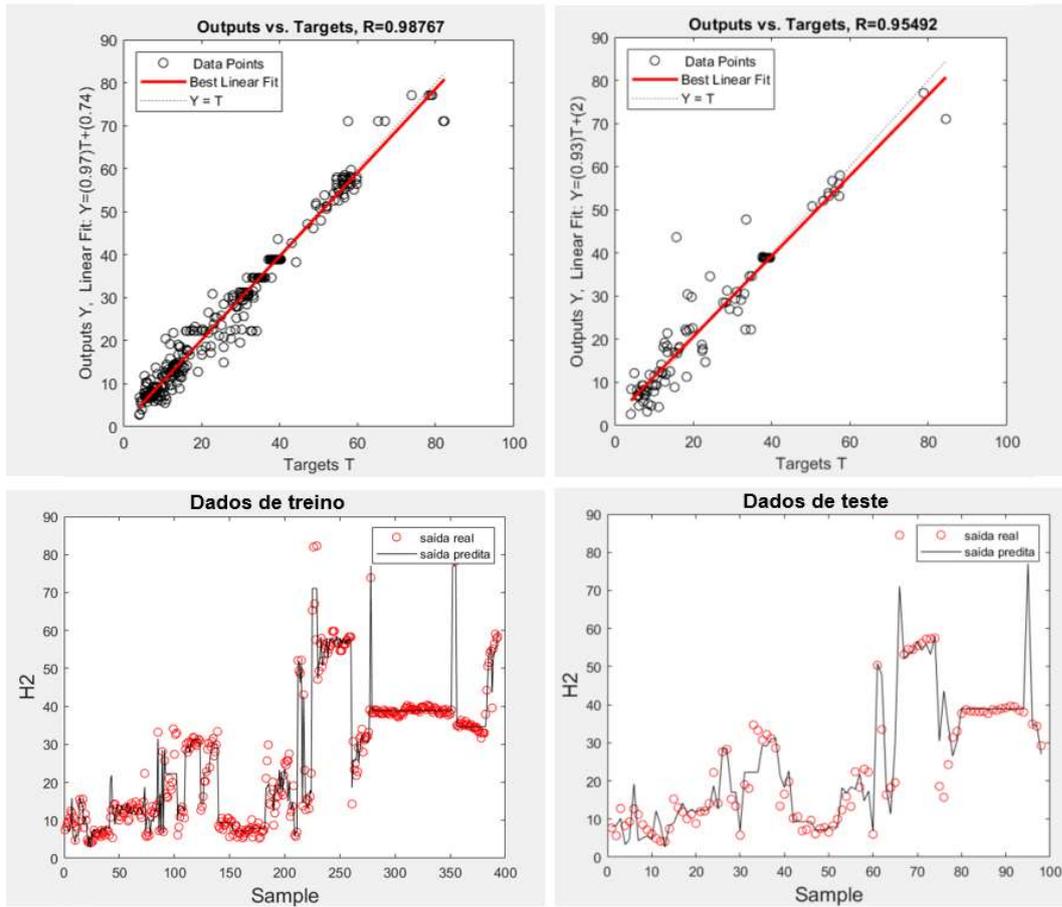


Figura 19 - Desempenho da melhor RNA para os dados de treinamento e teste – saída 1 (H<sub>2</sub>)

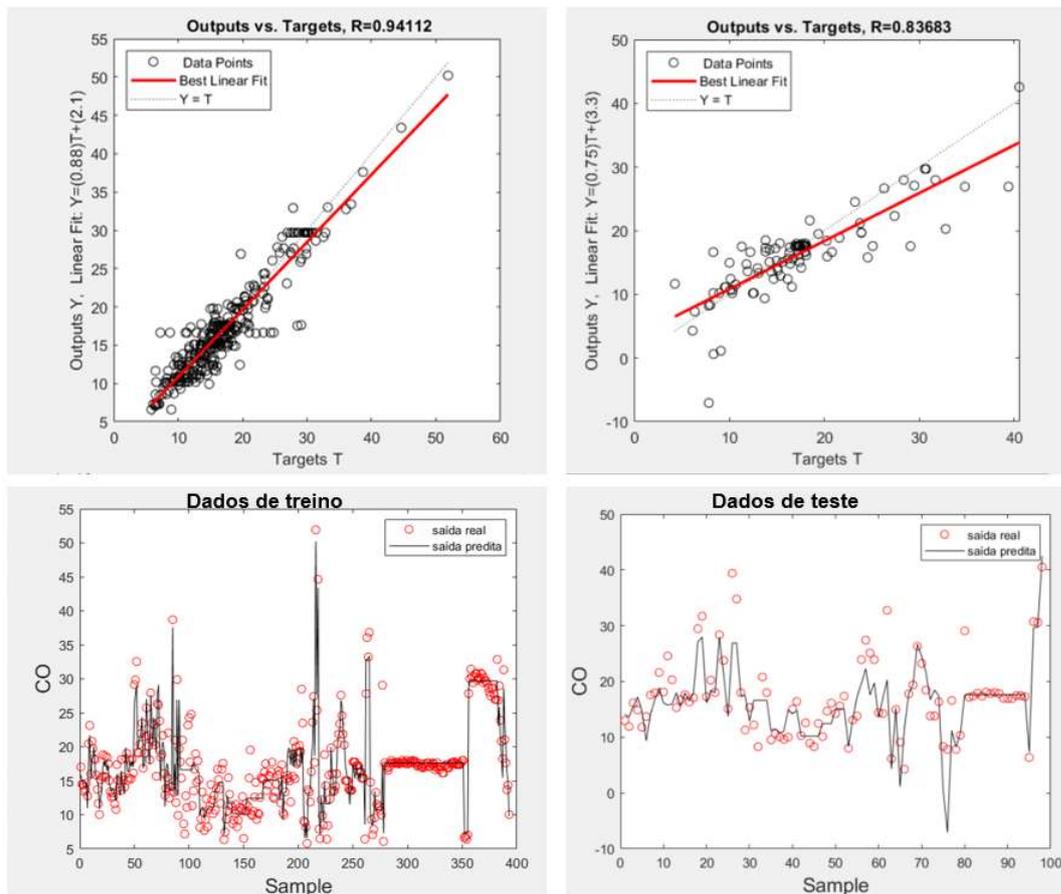


Figura 20 - Desempenho da melhor RNA para os dados de treino e teste – saída 2 (CO)

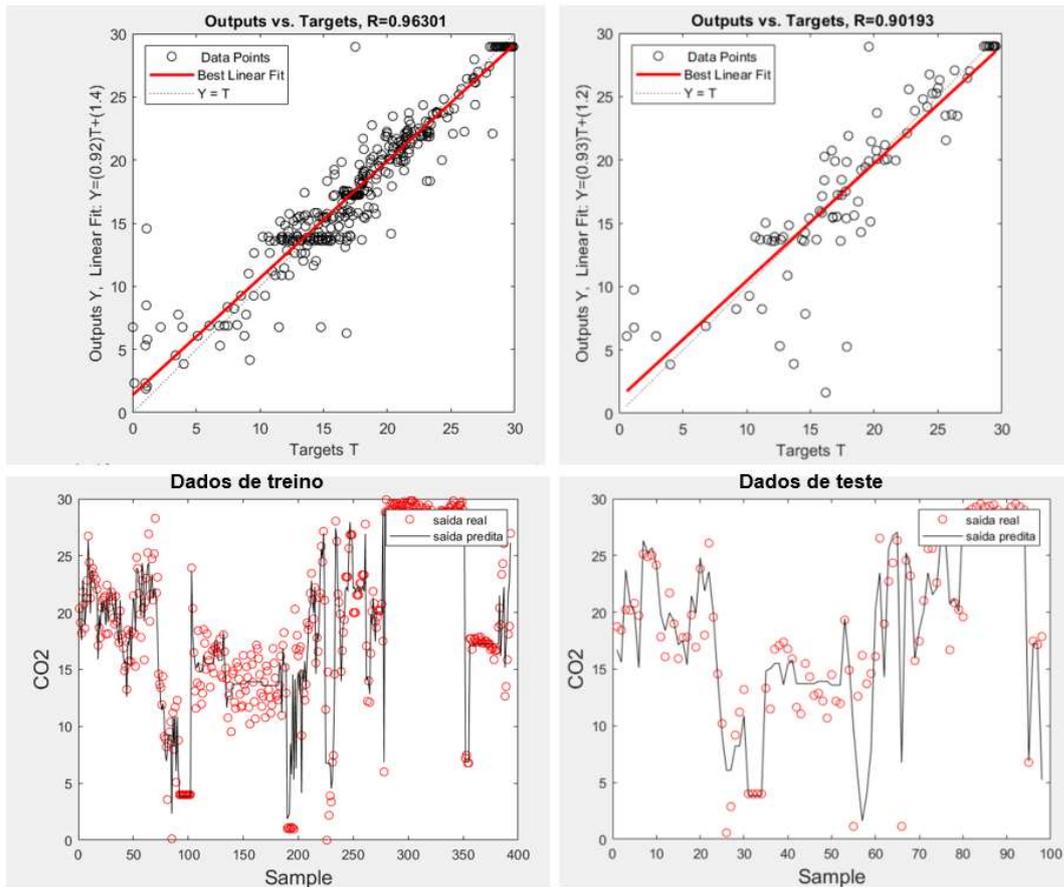


Figura 21 - Desempenho da melhor RNA para os dados de treino e teste – saída 3 (CO<sub>2</sub>)

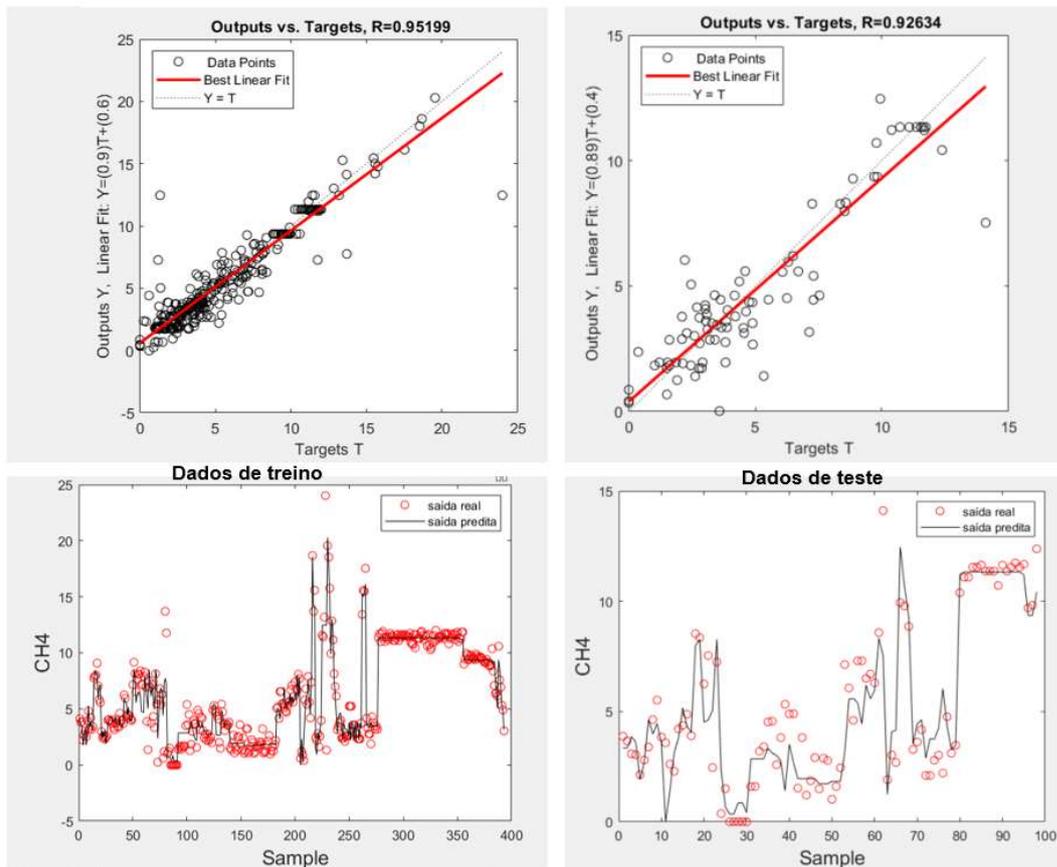


Figura 22 - Desempenho da melhor RNA para os dados de treino e teste – saída 4 (CH<sub>4</sub>)

Como pode ser observado nas Figuras 19, 21 e 22, os gráficos de *outputs vs. targets* indicam que os ajustes lineares para H<sub>2</sub>, CO<sub>2</sub> e CH<sub>4</sub> são próximos de  $Y=T$  e os gráficos H<sub>2</sub>, CO<sub>2</sub> ou CH<sub>4</sub> vs. *sample* indicam que a saída predita está próxima da saída real para a maioria dos dados tanto para o treinamento quanto para o teste, o que significa uma boa capacidade de predição para os referidos gases.

Além disso, pode-se afirmar que a capacidade de predição do modelo para o CO é um pouco inferior quando comparada aos demais gases. Isso se deve ao fato de que os ajustes lineares para o CO se afastam um pouco mais da reta  $Y=T$ , bem como uma quantidade maior de dados tem saída predita mais afastada da saída real (vide Figura 20). No entanto, ainda assim considera-se que a capacidade de predição do modelo para CO é satisfatória, tendo em vista que a média ponderada do R<sup>2</sup> de treino e teste é de aproximadamente 0,85.

### 4.3.2. Superfícies de resposta da RNA eleita

Para montagem das superfícies de resposta da RNA, foram escolhidos o teor de umidade (MC) e a temperatura, dentre as dez variáveis de entrada, para analisar como suas variações influenciam nos gases de saída. As demais variáveis de entrada – *i.e.*, carbono (C), hidrogênio (H), oxigênio (O), nitrogênio (N), material volátil (VM), cinzas, razão de equivalência (ER) e razão vapor/biomassa (S/B) – foram fixadas com seus valores médios de treinamento, que são, respectivamente: 48,99 %; 5,95 %; 44,39 %; 0,53 %; 75,51 %; 4,35 %; 0,19 %; e 0,39 %.

A escolha das duas variáveis de entrada para plotagem dos gráficos teve como fundamento os ensinamentos dos artigos Plis e Wilk (2011), Ebadi e Hisoriev (2017) e Sidek, *et. al.* (2020), os quais destacam que o teor de umidade e a temperatura influenciam significativamente na composição do gás de síntese, como pode ser visto nos trechos abaixo (tradução livre, ênfase adicionada):

*“O teor de umidade no combustível afeta muito os parâmetros de operação do gaseificador e a composição do gás de síntese.”* (Plis e Wilk, 2011, p. 3843).

*“A temperatura de gaseificação foi reconhecida como uma das variáveis operacionais mais efetivas, afetando a composição do gás de síntese produzido e o desempenho do gaseificador utilizado.”* (Ebadi e Hisoriev, 2017, resumo e item 3.2).

*“O parâmetro mais crucial na gaseificação da biomassa é a temperatura, que afetará a composição do gás, o teor de alcatrão e a taxa de reação.”* (Sidek, *et. al.*, 2020, p. 5)

De forma complementar, ressalta-se que os dados de treinamento relativos ao teor de umidade e à temperatura têm uma boa representatividade, uma vez que seus dados estão bem distribuídos e não há classes vazias, como pode ser visto nos histogramas abaixo.

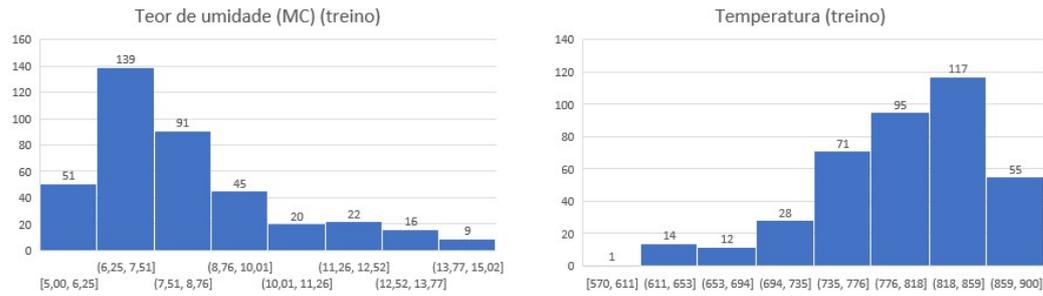


Figura 23 - Histogramas dos dados de treinamento do teor de umidade e temperatura

Não obstante, é válido comentar que, de acordo com a literatura, a razão de equivalência (ER) e razão vapor/biomassa (S/B) também são variáveis que influenciam significativamente na composição do gás de síntese. No entanto, como o modelo desenvolvido no presente trabalho engloba diversos tipos de agente de gaseificação e considera uma quantidade relevante de dados em que ER ou S/B é igual a zero, optou-se por não considerar tais variáveis de entrada na geração e análise das superfícies de resposta.

A Figura 24 contém os gráficos de superfície de resposta gerados para cada saída:

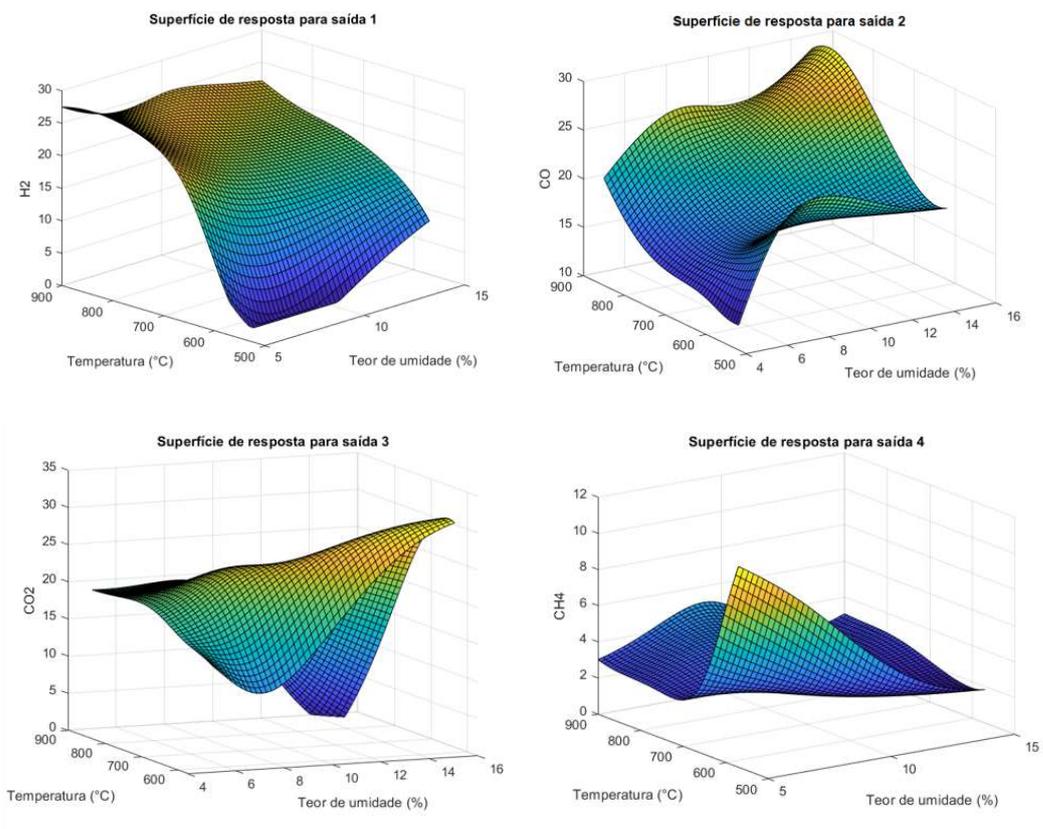


Figura 24 - Superfícies de resposta para as saídas: 1 - H<sub>2</sub>, 2 - CO, 3 - CO<sub>2</sub> e 4 - CH<sub>4</sub>

A partir da imagem da saída  $H_2$ , observa-se que quanto maiores os valores de temperatura, maior a produção de  $H_2$  e menor a influência do valor do teor de umidade. Do contrário, faixas de temperatura entre 550 e 600°C, por exemplo, além de produzirem menos  $H_2$ , essa produção é muito dependente do teor de umidade (quanto menor teor de umidade, menos  $H_2$  gerado para essa faixa de temperatura). Nesse contexto, é válido comentar que, de acordo com Sidek (2020), para a produção de *syngas* rico em hidrogênio, deve-se trabalhar com altas temperaturas, acima de 850°C. Esse comportamento é observado na imagem da saída  $H_2$  da Figura 24.

A partir da imagem da saída CO, observa-se que quanto maior a temperatura e maior o teor de umidade, maior a produção de CO e quanto mais baixos os valores de temperatura e do teor de umidade, menor a produção de CO. No estudo realizado por Zhou (2009), chegou-se à conclusão de que a temperatura teve uma forte influência no processo de gaseificação e que as temperaturas mais altas favoreceram a produção de  $H_2$  e de CO, ou seja, maiores temperaturas geram um *syngas* com mais  $H_2$  e CO. Esse comportamento é observado nas imagens das saídas  $H_2$  e CO da Figura 24.

A partir da imagem da saída  $CO_2$ , observa-se maior produção de  $CO_2$  para altos valores de teor de umidade e valores mais baixos de temperatura (entre 500 e 600°C). É observado também que para combinações de **(i)** temperatura alta e teor de umidade alto e **(ii)** temperatura baixa e teor de umidade baixo, há uma menor geração de  $CO_2$ .

A partir da imagem da saída  $CH_4$ , observa-se que quanto mais baixo o valor de temperatura e mais baixo o teor de umidade, mais elevada é a produção de  $CH_4$ . Nota-se também que em aproximadamente 8,7% de teor de umidade e 671°C de temperatura está a região de menor produção de  $CH_4$ . Ademais, ao trabalhar com uma temperatura de aproximadamente 900°C, para se obter maior produção de  $CH_4$ , deve-se trabalhar com uma biomassa com teor de umidade de cerca de 9%.

Como mencionado anteriormente, o *syngas* atua como matéria-prima para sínteses de biocombustíveis e produtos químicos e para geração de energia. Como exemplos da utilização do *syngas* pode-se listar a síntese de amônia, a produção de bio-hidrogênio e a geração de bio-metanol (Sikarwar et al., 2017). Nesse sentido, tendo em vista as superfícies de respostas apresentadas acima na Figura 24, as quais

consideram agentes de gaseificação mistos, são aqui tecidos alguns comentários analíticos:

- (i) Caso o interesse de aplicação da biomassa seja na produção de bio-hidrogênio, deve-se trabalhar com altas temperaturas de gaseificação, por exemplo, acima de 800°C;
- (ii) Caso o interesse de aplicação da biomassa seja na produção de bio-metanol, é interessante a obtenção de H<sub>2</sub>, CO e CO<sub>2</sub>, tendo em vista as principais reações de equilíbrio de síntese do bio-metanol (vide equações (1) e (2) do presente trabalho). Nesse contexto, é recomendável ter, por exemplo, uma temperatura de gaseificação de cerca de 860°C e teor de umidade de cerca de 10%, posto que nessas condições há boas produções de H<sub>2</sub>, CO e CO<sub>2</sub>, como pode ser visto na Figura 24;
- (iii) Caso o interesse de aplicação da biomassa seja na produção de combustível biometano, deve-se trabalhar com valores de temperatura de gaseificação próxima de 570°C e teor de umidade da biomassa próximo de 5%, como pode ser visto na Figura 24.

#### 4.4. Validação do modelo

Para validação do modelo de RNA desenvolvido (rede 6) foi utilizado o segundo banco de dados previamente preparado contendo 112 dados. Ressalta-se que os referidos dados não estão contidos no banco de dados principal, que foi utilizado para treinamento e teste da rede.

A seguir, são apresentados os gráficos das saídas (“*outputs*”) versus o alvo (“*target*”), bem como de saídas real e predita para cada amostra de dado ao executar a rede 6 para os dados de validação para H<sub>2</sub>, CO, CO<sub>2</sub> e CH<sub>4</sub> (Figuras 25 a 28 abaixo). Nos gráficos do lado esquerdo das Figuras abaixo, o alvo ou *target T* representa o H<sub>2</sub>, CO, CO<sub>2</sub> ou CH<sub>4</sub> do banco de dados, enquanto a saída (“*Output Y*”) é a predição feita pela RNA do alvo T. O ajuste linear é a regressão linear entre *Output Y* e *Target T* e sua forma ideal seria  $Y = T$ .

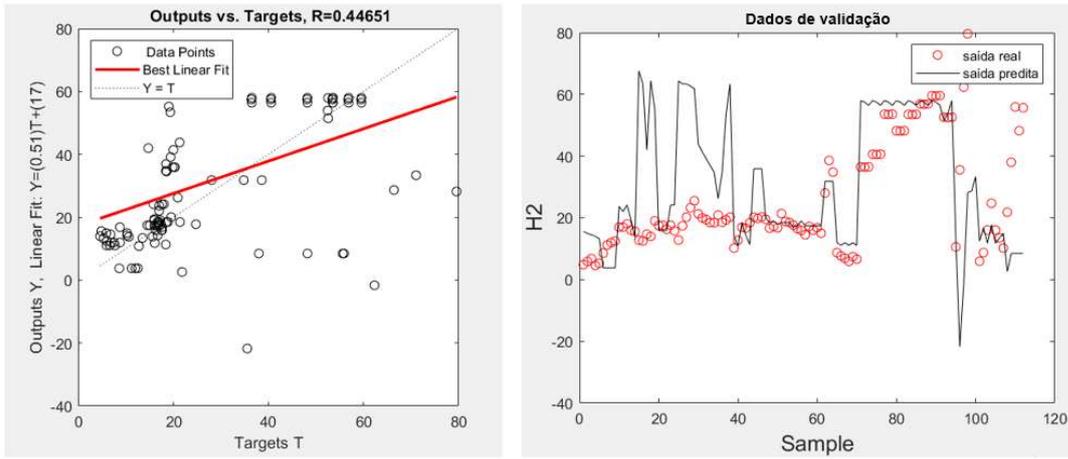


Figura 25 - Desempenho da melhor RNA para os dados de validação – saída 1 (H<sub>2</sub>)

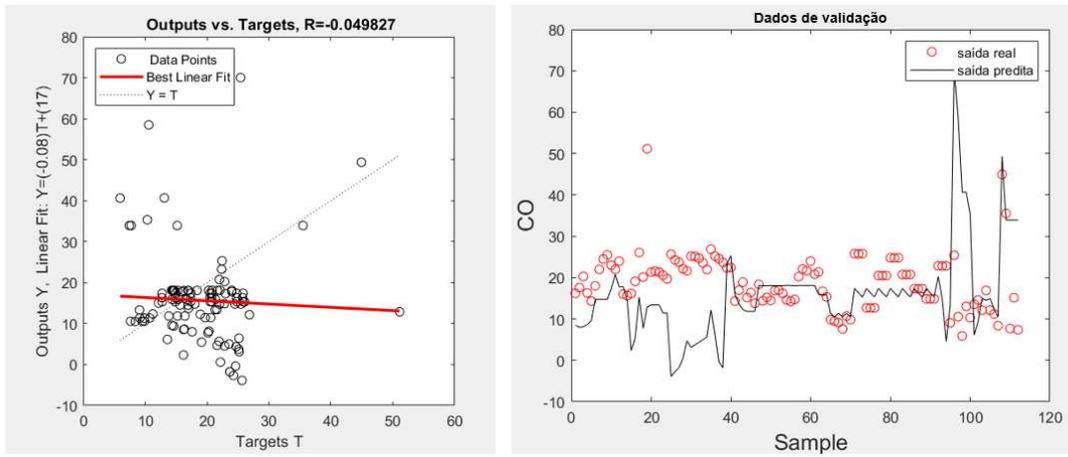


Figura 26 - Desempenho da melhor RNA para os dados de validação – saída 2 (CO)

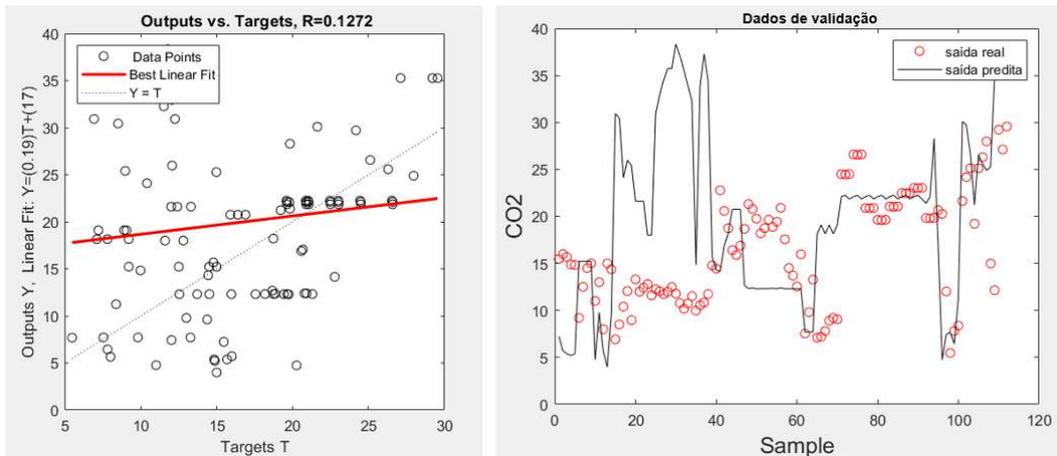


Figura 27 - Desempenho da melhor RNA para os dados de validação – saída 3 (CO<sub>2</sub>)

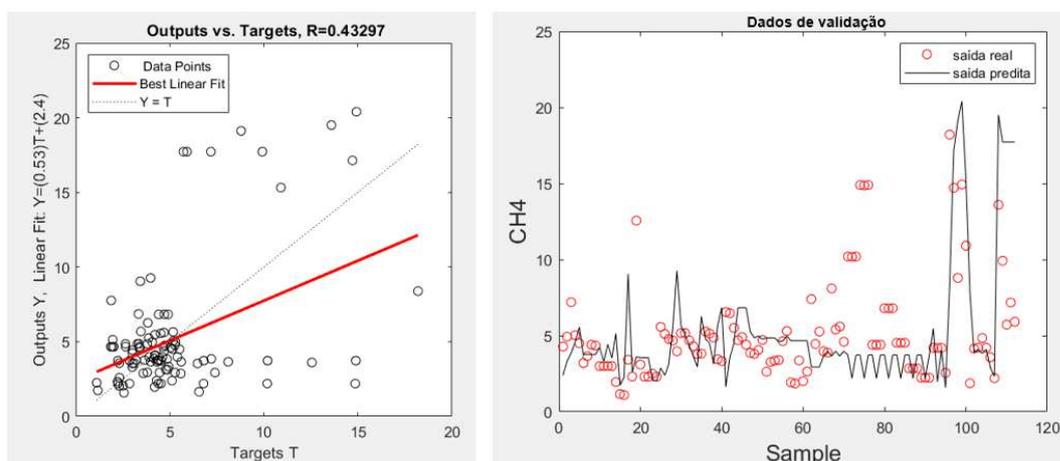


Figura 28 - Desempenho da melhor RNA para os dados de validação – saída 4 (CH<sub>4</sub>)

Subsequentemente, é apresentada uma tabela com os valores de  $R^2$  e SSE de cada uma das saídas da validação.

Tabela 6 - Valores de  $R^2$  e SSE de cada uma das saídas da validação

	Saída 1 (H <sub>2</sub> )	Saída 2 (CO)	Saída 3 (CO <sub>2</sub> )	Saída 4 (CH <sub>4</sub> )
$R^2$	0,1994	0,0025	0,0162	0,1875
SSE	$2,42 \times 10^5$	$1,34 \times 10^5$	$1,54 \times 10^5$	2,9994

Como pode ser observado nas Figuras 25 a 28 e Tabela 10, os resultados obtidos são considerados não satisfatórios. Os gráficos de *outputs vs. targets* indicam que os ajustes lineares para H<sub>2</sub>, CO, CO<sub>2</sub> e CH<sub>4</sub> não são próximos de  $Y=T$  e os gráficos H<sub>2</sub>, CO, CO<sub>2</sub> ou CH<sub>4</sub> vs. *sample* indicam que a saída predita está afastada da saída real para muitos dados, o que significa uma capacidade ruim de predição para os referidos gases. Novamente, o pior comportamento observado é para o CO, cujo  $R^2$  de validação é próximo de zero. Diante desse resultado, fez-se necessária uma reavaliação do comportamento dos dados que pudessem justificar valores de  $R^2$  tão baixos, quando a rede 6 tinha bons valores de  $R^2$  de treinamento e de teste. Tal avaliação está detalhada a seguir.

#### 4.4.1. Análises qualitativa e quantitativa dos dados e perspectivas para melhoria dos modelos

##### 4.4.1.1. Análise qualitativa

Como já comentado, ao criar o banco de dados de validação, foi feita uma análise cuidadosa para certificar que os dados de validação estavam compreendidos nas faixas de valores dos dados de treinamento, para as variáveis de entrada e de saída, de forma a não haver extrapolação dos primeiros. No entanto, não foi realizada uma verificação de como os dados de validação das variáveis de saída estavam distribuídos em relação aos dados de saída do treinamento, como foi feito com os dados do teste (vide item 4.1. Divisão dos dados de treinamento e de teste).

Fazendo esta análise, notou-se que os dados de validação para os gases de saída não estavam bem balanceados em relação aos dados de treinamento, principalmente com relação ao H<sub>2</sub>, como pode ser visto nas Figura 29 a 32 abaixo, o que pode ter contribuído para o resultado insatisfatório.

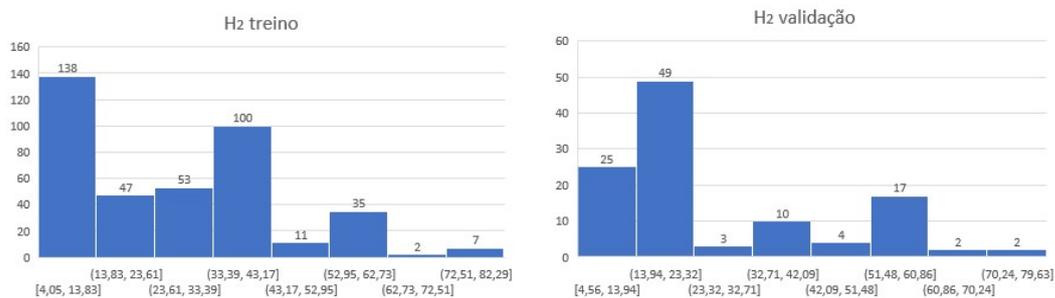


Figura 29 - Histogramas dos dados de treino e validação para o H<sub>2</sub>

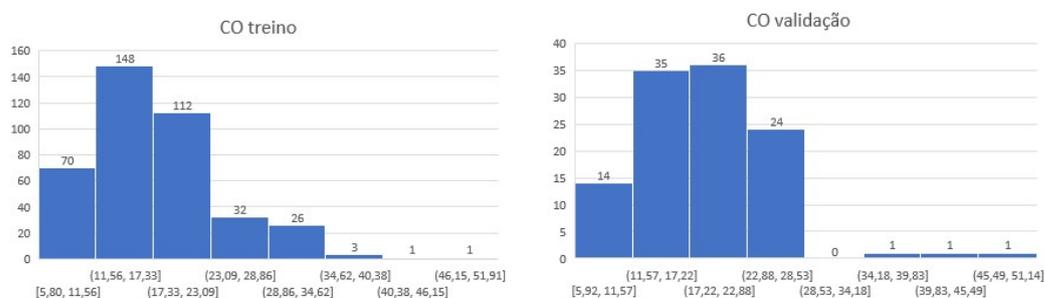


Figura 30 - Histogramas dos dados de treino e validação para o CO

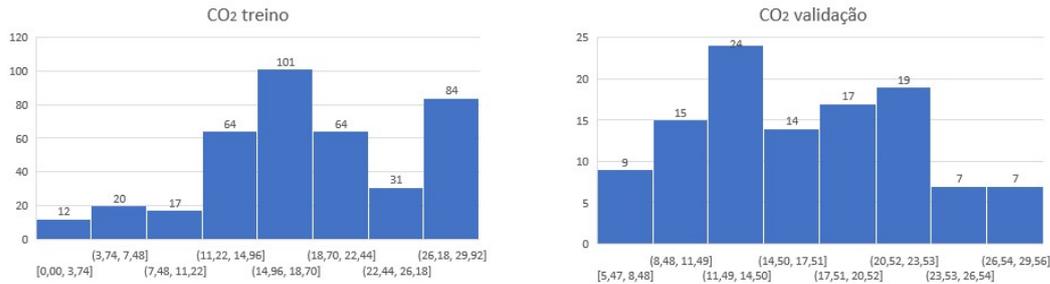


Figura 31 - Histogramas dos dados de treino e validação para o CO<sub>2</sub>

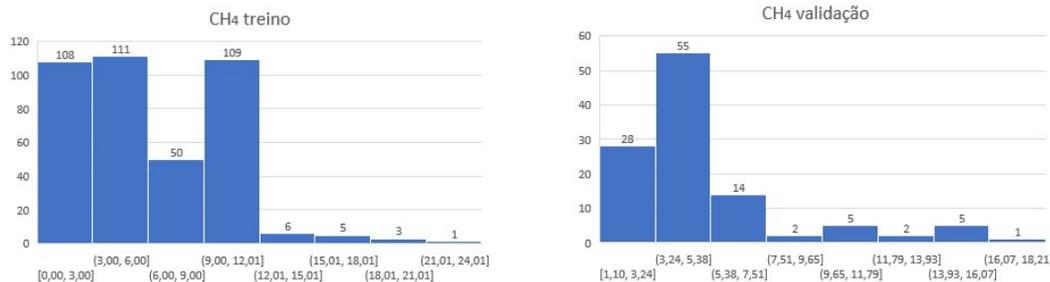


Figura 32 - Histogramas dos dados de treino e validação para o CH<sub>4</sub>

Destaca-se que, com base na distribuição dos dados do treinamento do H<sub>2</sub> (vide Figura 29), a rede foi bem treinada principalmente na faixa de valores ente 4,05 – 13,83% e 33,39 – 43,17%. Todavia, os dados de validação estão mais concentrados na faixa entre 13,94 e 23,32%.

Além disso, observando a distribuição dos dados das variáveis de entrada da validação em relação aos dados de treinamento, é possível notar que há regiões que foram pouco ou não foram treinadas (classes vazias ou relativamente pouco populosas nos histogramas dos dados de treino), mas que são populosas para os dados de validação. Quando há muitos dados de validação em regiões que não foram bem treinadas, é muito provável que o modelo tenha um desempenho ruim. Ao contrário, nas regiões bem treinadas (onde ocorreu o aprendizado), o modelo funciona melhor. Para pronta visualização das distribuições dos dados de forma comparativas entre treino e validação, os histogramas respectivos (contidos no **Apêndice IV**) serão aqui agrupados e comentados.

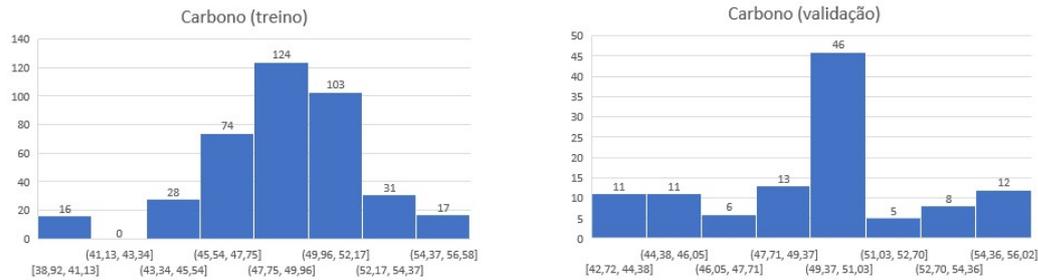


Figura 33 - Histogramas dos dados de treino e validação para o Carbono

Como pode ser observado a partir da Figura 33, no banco de dados de validação, existem 6 dados para a classe vazia observada no histograma do treinamento, relativa à faixa de 41,13 – 43,34% de Carbono.

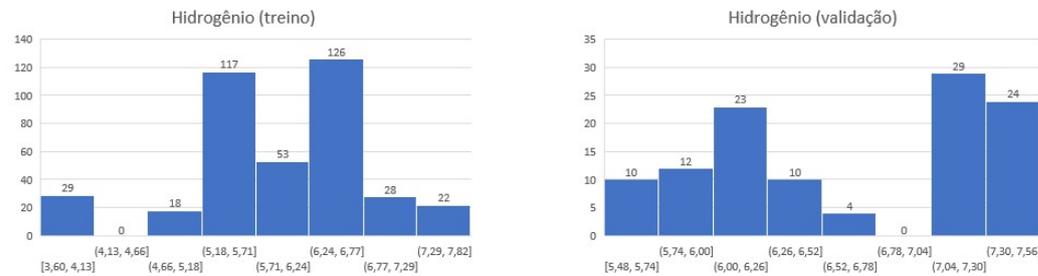


Figura 34 - Histogramas dos dados de treino e validação para o Hidrogênio

Como pode ser observado a partir da Figura 34, as classes mais populosas do treinamento do hidrogênio – e, portanto, as regiões de maior aprendizagem – se referem às classes dos intervalos de 5,18-5,71% e 6,24-6,77%. Para esses dois intervalos de valores, existem, respectivamente, apenas 10 e 14 dados no grupo de dados de validação. Ressalta-se ainda que as regiões onde a fração mássica de H é maior a 6,77% tem relativamente poucos eventos para treinamento (duas últimas classes do histograma da esquerda da Figura 34 acima), são as regiões mais populosas da validação.

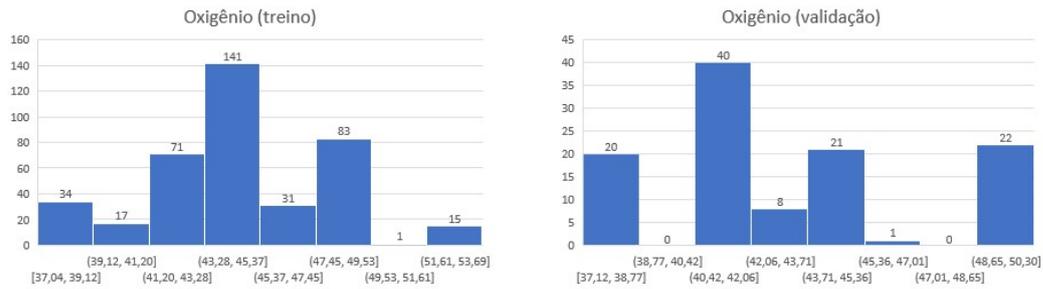


Figura 35 - Histogramas dos dados de treino e validação para o Oxigênio

Como pode ser observado a partir da Figura 35, no banco de dados de validação, existem 6 dados para a classe do histograma de treinamento que compreende apenas um evento, *i.e.*, relativa à faixa de 49,53 – 51,61% de Oxigênio.

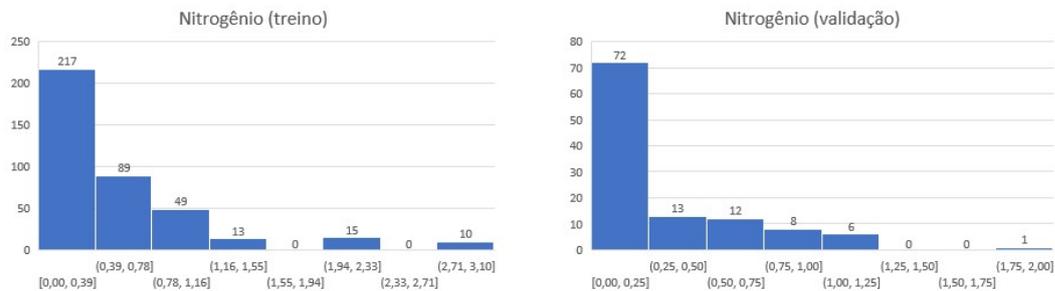


Figura 36 - Histogramas dos dados de treino e validação para o Nitrogênio

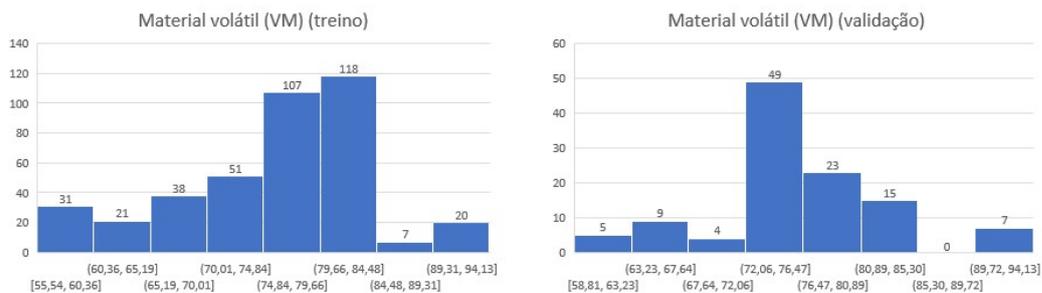


Figura 37 - Histogramas dos dados de treino e validação para o material volátil (VM)

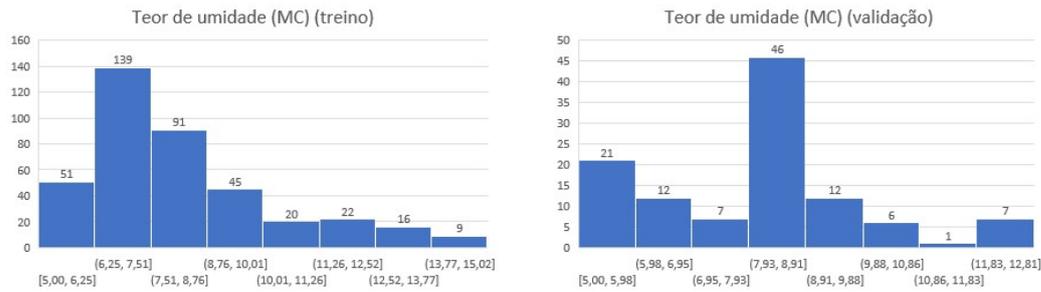


Figura 38 - Histogramas dos dados de treino e validação para o teor de umidade (MC)

As regiões mais populosas da validação para Nitrogênio (0 - 0,25% de N), material volátil (72,06 – 85,30 % de VM) e teor de umidade (7,93 – 8,91% de MC) foram bem treinadas, como pode ser visto nas Figuras 36 a 38. Além disso, com relação ao Nitrogênio (Figura 36), os dados de validação estão bem distribuídos, tendo em vista os dados de treinamento. Já com relação ao VM e MC (Figuras 37 e 38), as distribuições dos dados de validação não estão tão semelhante às distribuições dos dados de treinamento.

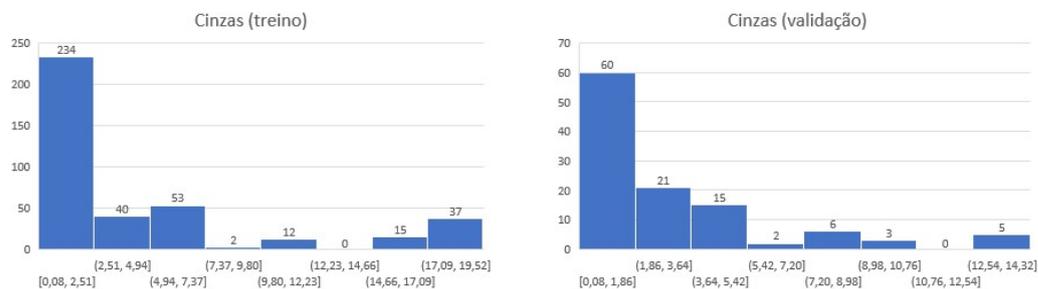


Figura 39 - Histogramas dos dados de treino e validação para cinzas

Como pode ser observado na Figura 39, no banco de dados de validação, existem 5 dados para a classe vazia observada no histograma do treinamento, relativa à faixa de 12,23 – 14,66% de cinzas. Além disso, é válido comentar que no banco de dados de validação existem 5 dados para a classe do histograma de treinamento que compreende apenas dois eventos (baixa representatividade com relação ao número de dados de treinamento), *i.e.*, relativa à faixa de 7,37 – 9,80% de cinzas.

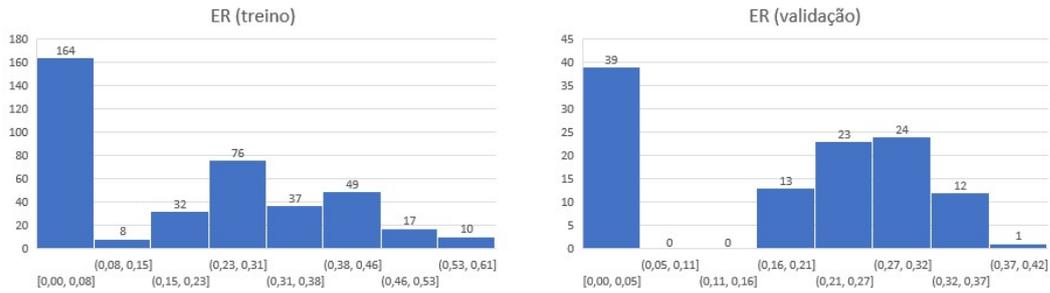


Figura 40 - Histogramas dos dados de treino e validação para ER

Com relação ao ER, como pode ser observado na Figura 40, os dados de validação estão bem distribuídos, tendo em vista os dados de treinamento. As regiões mais populosas da validação (0 – 0,05% e 0,21 – 0,32 de ER) foram bem treinadas.

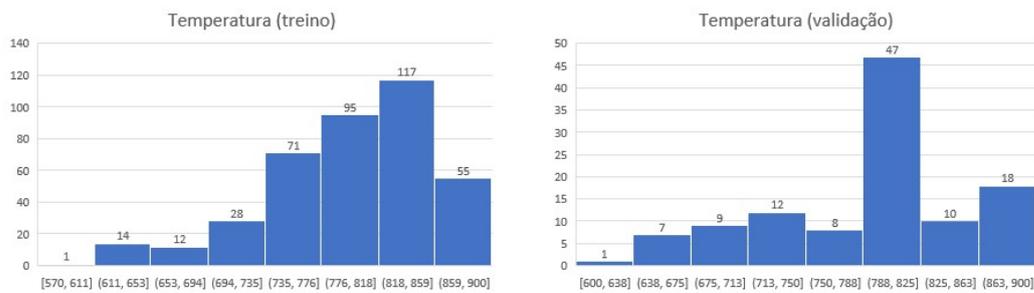


Figura 41 - Histogramas dos dados de treino e validação para temperatura

Como pode ser observado na Figura 41, apesar do perfil da distribuição dos dados de validação não estar tão semelhante à distribuição dos dados de treinamento para temperatura, a região mais populosa da validação (788 – 825°C de T) foi bem treinada.

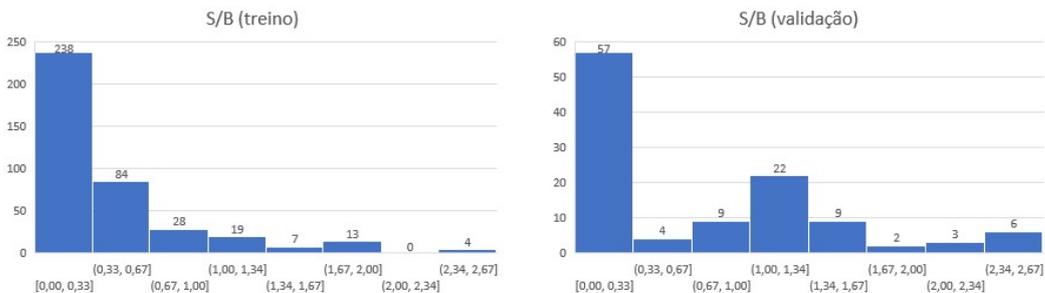


Figura 42 - Histogramas dos dados de treino e validação para S/B

Como pode ser observado a partir da Figura 42, no banco de dados de validação, existem 3 dados para a classe vazia observada no histograma do treinamento, relativa à faixa de valores entre 2,00 e 2,34 de S/B.

Diante da análise qualitativa acima, conclui-se que o desbalanceamento na distribuição dos dados das variáveis de entrada, principalmente C, H, O, cinzas e S/B, pode ser um dos motivos para o desempenho ruim obtido na validação do modelo da RNA. A RNA torna-se especialista nas regiões que são mais populosas para os dados de treinamento, nas quais o aprendizado é maior, e defasada em outras regiões, onde há poucos eventos de treinamento. No entanto, quando há dados de validação nessas últimas regiões, o desempenho da RNA na validação fica comprometido.

Uma possível solução para o desempenho baixo obtido na validação seria retrabalhar o banco de dados de treinamento no sentido de incluir mais dados em regiões que atualmente existem classes vazias.

#### 4.4.1.2. Análise quantitativa

O objetivo da presente análise é avaliar quantitativamente o comportamento dos dados de validação em relação aos dados de treinamento. Para tal, foi feita uma análise das distâncias das variáveis de entrada dos dados de validação em relação aos dados de treinamento em um hiperespaço em função do desvio dos gases de saída (métrica de erro adotada). As equações a seguir representam os cálculos das distâncias de um ponto de treinamento e um ponto de validação (23) e do desvio dos gases de saída (24):

$$D(i,j) = (\sum_{k=1}^{10} (X_{\text{treinamento}}(i,k) - X_{\text{validação}}(j,k))^2)^{1/2} \quad (23)$$

em que:

$D(i,j)$  é a distância entre um ponto do banco de treinamento  $i$  ( $i$  varia entre 1 e 393) e um ponto do banco de validação  $j$  ( $j$  varia entre 1 e 112);

$X_{\text{treinamento}}(i,k)$  e  $X_{\text{validação}}(j,k)$  representam a  $k$  variável de entrada ( $k$  varia entre 0 e 10) para o ponto do banco de treinamento  $i$  e a  $k$  variável de entrada ( $k$  varia entre 0 e 10) para o ponto do banco de validação  $j$ , respectivamente.

$$\text{Desvio} = ((\text{H}_2 \text{ BD} - \text{H}_2 \text{ RNA})^2 + (\text{CO} \text{ BD} - \text{CO} \text{ RNA})^2 + (\text{CO}_2 \text{ BD} - \text{CO}_2 \text{ RNA})^2 + (\text{CH}_4 \text{ BD} - \text{CH}_4 \text{ RNA})^2)^{1/2} \quad (24)$$

em que o índice BD é o valor de saída do banco de dados de validação e o índice RNA é o valor de saída calculado pelo modelo de rede neural artificial.

Para a análise das distâncias, inicialmente foi gerada uma matriz, em que cada coluna corresponde a um vetor com as variáveis de entrada de um ponto do banco de validação e cada linha correspondente a um vetor com as variáveis de entrada de um ponto do banco de treinamento, conforme ilustrado na Tabela 11. As distâncias foram calculadas com base na equação (23).

Tabela 7 – Ilustração da matriz distância de treinamento e validação

Pontos	P1 validação	P2 validação	...	P112 validação
<b>P1 treino</b>	Distância entre os vetores	Distância entre os vetores		Distância entre os vetores
<b>P2 treino</b>	Distância entre os vetores	Distância entre os vetores		Distância entre os vetores
...				
<b>P393 treino</b>	Distância entre os vetores	Distância entre os vetores		Distância entre os vetores

Em seguida, as colunas da matriz com as distâncias foram ordenadas de forma crescente para pronta visualização das menores distâncias, já que menores valores significam pontos mais próximos da validação em relação ao treinamento. A matriz foi feita no Matlab R2022a, e o código está contido no **Apêndice V** (denominado “Matriz para análise quantitativa”).

Inicialmente foram plotados os gráficos da 1ª e 12ª menores distâncias (Figura 43 abaixo). No entanto, observou-se uma alteração da localização de uma aglomeração dos pontos e que, quanto maior o desvio (eixo x), maior é o valor da distância (eixo y) desse desvio.

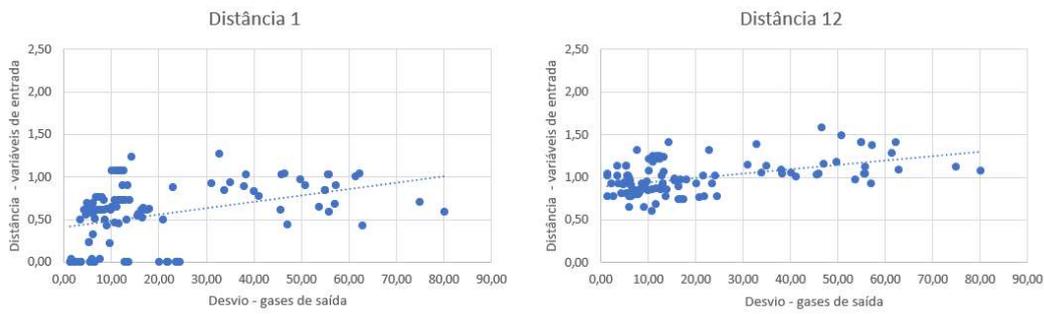
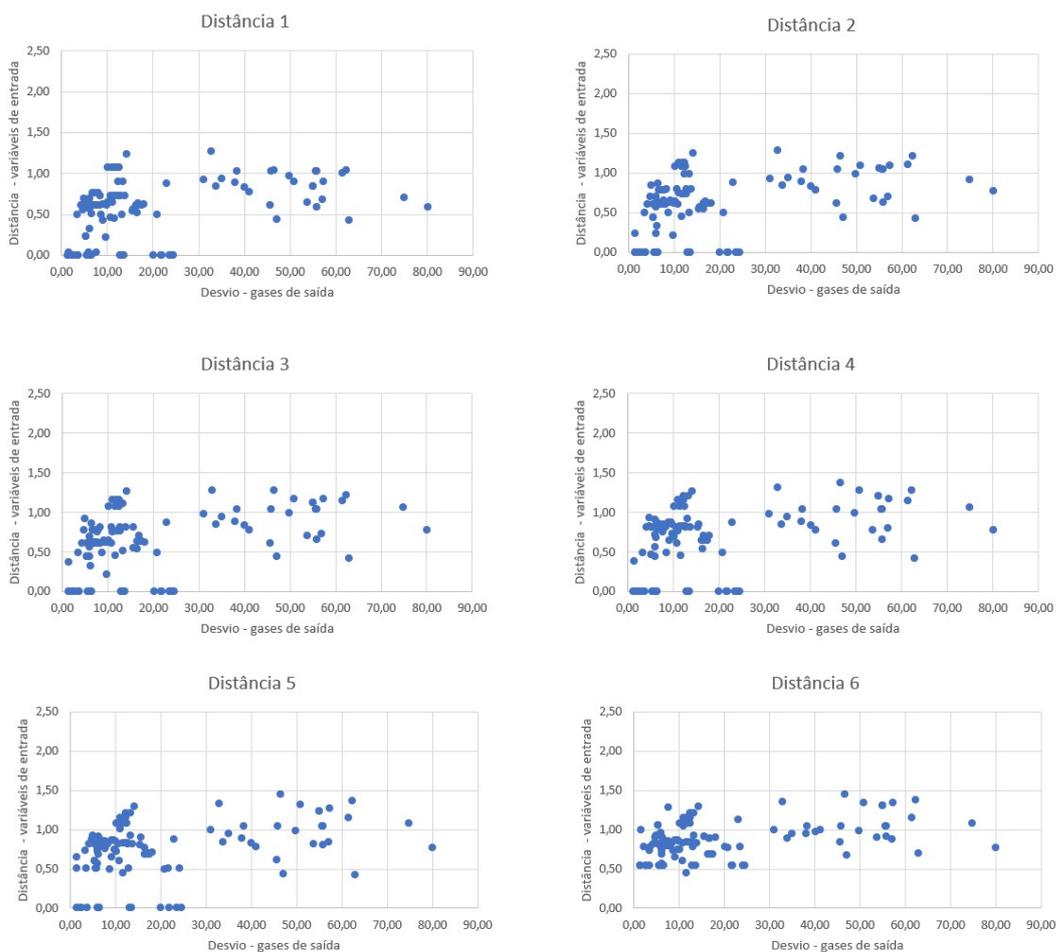


Figura 43 - Gráficos distância x desvio das 1ª e 12ª menores distâncias

Dessa maneira, foi considerado ser interessante acompanhar a evolução do comportamento dos pontos entre a 1ª e 12ª menores distâncias, bem como checar o comportamento das 15ª, 20ª e 50ª menores distâncias. As figuras a seguir correspondem aos gráficos das 1ª a 12ª, 15ª, 20ª e 50ª menores distâncias.



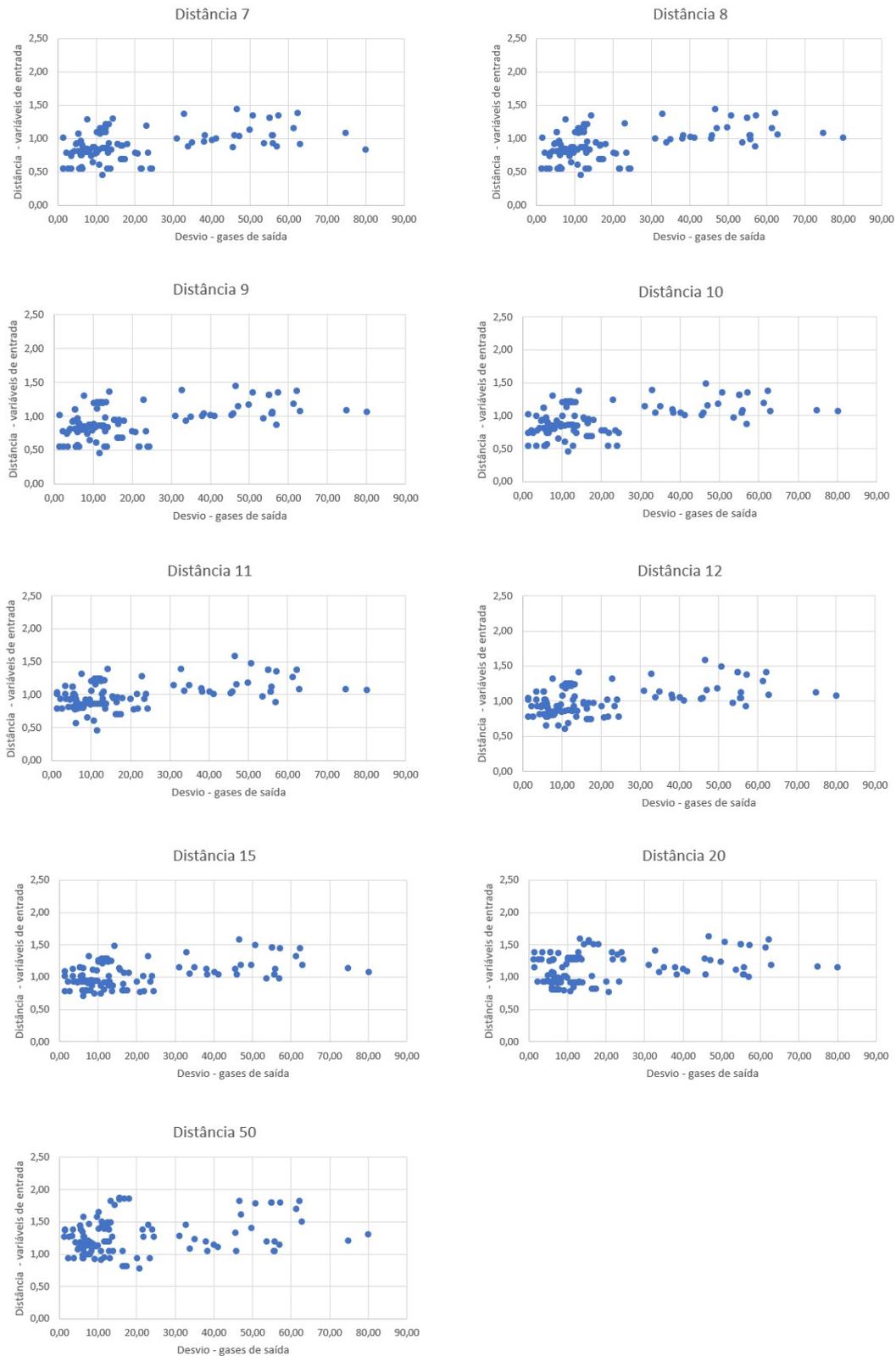


Figura 44 - Gráficos distância x desvio das 1ª a 12ª, 15ª, 20ª e 50ª menores distâncias

Primeiramente, destaca-se que o melhor é ter: (1) menores distâncias (pontos mais para a parte inferior em relação ao eixo y), pois indica que os pontos de

validação estão mais próximos dos pontos de treinamento, onde houve aprendizado; e (2) menores desvios (pontos mais para esquerda, em relação ao eixo x), pois indica que as saídas previstas estão mais próximas das saídas observadas (do banco de dados/literatura).

Como comentário geral, observa-se pelos gráficos da Figura 44 acima que em todos eles existem alguns pontos dispersos mais para direita do gráfico e há uma aglomeração até mais ou menos o valor 25 de desvio. O número considerável de dados dispersos à direita dos gráficos, *i.e.*, erro acumulado maior que 25%, pode ter contribuído para o desempenho ruim da validação.

Considera-se que o gráfico da Distância 7 – ou seja, a sétima menor distância – pode ser assumida como a distância de corte, assumindo que erros acumulados até 25% são aceitáveis. A partir do gráfico da Distância 6, ocorre uma mudança significativa na posição da nuvem de pontos em relação ao eixo y e isso fica mais claro no gráfico da Distância 7.

Observando os pontos do gráfico da Distância 7 da Figura 44, pode-se assumir que as menores distâncias aceitáveis entre dados de validação e de treinamento são de até 0,6448.

Diante do acima, uma possível solução para o desempenho baixo obtido na validação seria retrabalhar o banco de dados de validação no sentido de incluir mais dados mais próximos de regiões bem treinadas.

É fundamental destacar que estamos tratando, no presente trabalho, de banco de dados amplos, os quais englobam diferentes tipos de biomassa, de agente de gaseificação e de gaseificadores e essa diversidade de dados também é um dificultador para a robustez do modelo. Tendo em vista as análises qualitativa e quantitativa realizadas, como perspectivas futuras, a sugestão é aplicar técnicas de balanceamento de dados antes do fechamento dos bancos de dados e desenvolvimento dos modelos de RNA, por exemplo, reestruturando os dados e/ou coletando mais dados.

#### **4.5. Avaliação da robustez na região de validação**

Apesar da validação do modelo não ter sido muito adequada para o conjunto de dados em questão, conforme descrito na seção anterior, foi possível identificar

com uma métrica quantitativa simples as regiões mais confiáveis onde há uma maior densidade de dados no treinamento.

A partir de pontos característicos das superfícies de resposta apresentadas na seção 4.3.2 do presente trabalho, ou seja, os pontos que correspondem aos vértices do domínio investigado e um ponto central, foi gerada uma nova matriz, em que cada coluna corresponde a um vetor com as variáveis de entrada dos pontos característicos e cada linha correspondente a uma das variáveis de entrada de um ponto do banco de treinamento, de forma similar à Tabela 11 contida na seção 4.4.1.2 do presente trabalho, mas com uma dimensão de 393 linhas (pontos de treino) e 5 colunas (P1 com MC = 5, T = 570, demais = média; P2 com MC = 5, T = 900, demais = média; P3 com MC = 15,02, T = 570, demais = média; P4 com MC = 15,02, T = 900, demais = média; P5 com MC = 9,75, T = 724,898, demais = média).

Novamente, as colunas da matriz com as distâncias foram ordenadas de forma crescente. A 7ª menor distância para cada ponto característico foi incluída no gráfico previamente plotado para a 7ª menor distância entre os dados de validação e de treinamento, como pode ser visto a seguir:

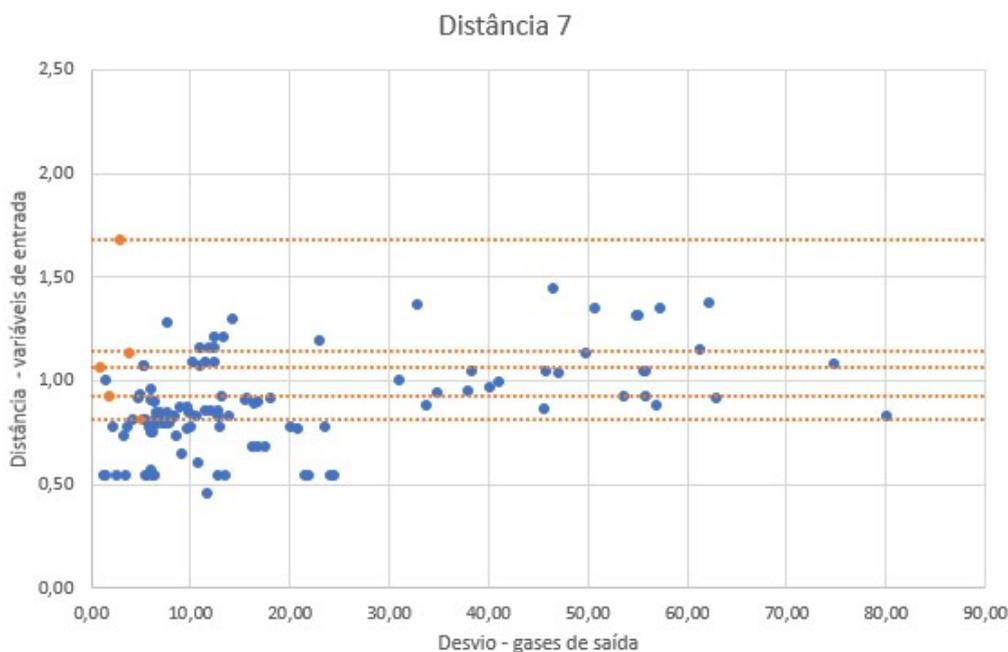


Figura 45 - Gráfico distâncias 7

A partir da figura acima, foi possível confirmar que é ideal gerar curvas de superfície para um subgrupo de validação. A ideia é selecionar um subgrupo bem

representativo dos dados de treinamento, usando o critério definido das distâncias (até distância 7). Como mencionado na seção anterior, foi assumido que as menores distâncias aceitáveis entre dados de validação e de treinamento são de até 0,6448. Esse valor e essa distância foram escolhidas de maneira a reduzir a quantidade de condições experimentais que poderia ser considerada de maneira incorreta como tendo um desempenho satisfatório. Em outras palavras, com essa escolha, pretende-se reduzir a quantidade de falsos positivos. A quantidade de falsos negativos (condições experimentais que tem um desempenho satisfatório são consideradas de maneira incorreta) pode ser significativa e foi usada como critério de decisão para a escolha das distâncias 7. Identificou-se então todos os casos na sétima distância, *i.e.*, na linha 7 da matriz distância treinamento x validação, que têm distâncias menores que 0,6448. Em seguida, foi identificado quais pontos da validação se referem esses casos e assim selecionando, no banco de dados de validação, esses pontos de validação que equivalem às menores distâncias na linha 7 identificados na matriz, resultando em um subgrupo com 18 dados de validação, conforme ilustrado abaixo:

	C [%wt db]	H [%wt db]	O [%wt db]	N [%wt db]	Volatile matter (VM) [%wt]	Moisture content (MC) [%wt]	Ash [%wt db]	ER [-]	T [°C]	Steam/Biomass	H <sub>2</sub> [%vol]	CO [%vol]	CO <sub>2</sub> [%vol]	CH <sub>4</sub> [%vol]
10	48,71	6,08	44,60	0,51	76,33	6,43	1,18	0,28	820	0,00	17,00	23,00	11,00	3,00
11	48,91	5,80	45,11	0,18	72,97	9,50	1,90	0,30	850	0,00	17,00	22,00	13,00	3,00
12	48,71	6,08	44,60	0,51	76,33	6,43	1,18	0,26	823	0,00	18,00	24,00	8,00	3,00
13	48,71	6,08	44,60	0,51	76,33	6,43	1,18	0,35	850	0,00	16,00	16,00	15,00	3,00
71	55,00	7,36	37,12	0,00	73,74	5,73	2,21	0,00	800	1,33	36,50	25,80	24,50	10,20
72	53,10	7,47	38,32	0,44	75,99	6,56	5,33	0,00	800	1,33	36,42	25,74	24,45	10,18
74	55,00	7,36	37,12	0,00	73,74	5,73	2,21	0,00	800	1,33	40,60	12,70	26,60	14,90
75	53,10	7,47	38,32	0,44	75,99	6,56	5,33	0,00	800	1,33	40,52	12,68	26,55	14,87
77	55,00	7,36	37,12	0,00	73,74	5,73	2,21	0,00	800	1,33	53,60	20,50	20,90	4,40
78	53,10	7,47	38,32	0,44	75,99	6,56	5,33	0,00	800	1,33	53,51	20,47	20,87	4,39
80	55,00	7,36	37,12	0,00	73,74	5,73	2,21	0,00	800	1,33	48,23	24,79	19,64	6,80
81	53,10	7,47	38,32	0,44	75,99	6,56	5,33	0,00	800	1,33	48,14	24,74	19,60	6,79
83	55,00	7,36	37,12	0,00	73,74	5,73	2,21	0,00	800	1,33	53,54	20,77	21,06	4,53
84	53,10	7,47	38,32	0,44	75,99	6,56	5,33	0,00	800	1,33	53,45	20,74	21,03	4,52
86	55,00	7,36	37,12	0,00	73,74	5,73	2,21	0,00	800	1,33	56,90	17,34	22,50	2,85
87	53,10	7,47	38,32	0,44	75,99	6,56	5,33	0,00	800	1,33	56,81	17,31	22,47	2,85
89	55,00	7,36	37,12	0,00	73,74	5,73	2,21	0,00	800	1,33	59,60	14,87	23,04	2,23
90	53,10	7,47	38,32	0,44	75,99	6,56	5,33	0,00	800	1,33	59,51	14,85	23,01	2,23

Figura 46 - Subgrupo de validação

Nota-se que os 4 primeiros dados (linhas 10 a 13) se referem a gaseificação com ar, enquanto os 14 restantes correspondem a casos de gaseificação com vapor e se referem a duas biomassas próximas. Foi então selecionada uma das duas biomassas em que o agente é vapor para fixar as variáveis de entrada para geração das superfícies de resposta, exceto T e MC, que irão variar.

No entanto, antes da geração das superfícies, foi feita uma verificação e ajuste se os quatro vértices e o centro do domínio investigado atenderiam ao critério de distância. Os novos pontos estão representados com as linhas em verde na Figura 47 abaixo. Todos eles estão em uma região melhor, mais bem treinada, quando comparados aos pontos em laranja, os quais foram plotados como ponto de partida com base nas extremidades e centro das superfícies de respostas dos dados de treinamento.

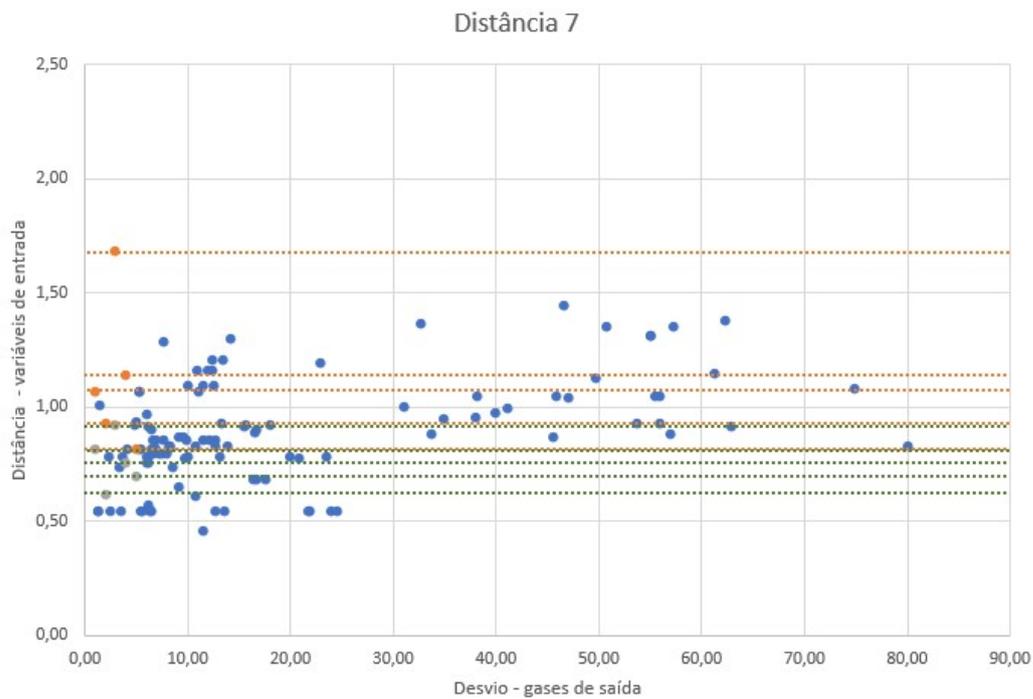


Figura 47 - Distância 7

Para geração das superfícies de resposta, foi selecionado um subgrupo dos dados de treinamento bem representativos, de acordo com a análise das distâncias acima mencionadas. As novas superfícies de resposta plotadas estão apresentadas abaixo, em que o MC varia de 5 a 10 e a temperatura de 700 a 850°C.

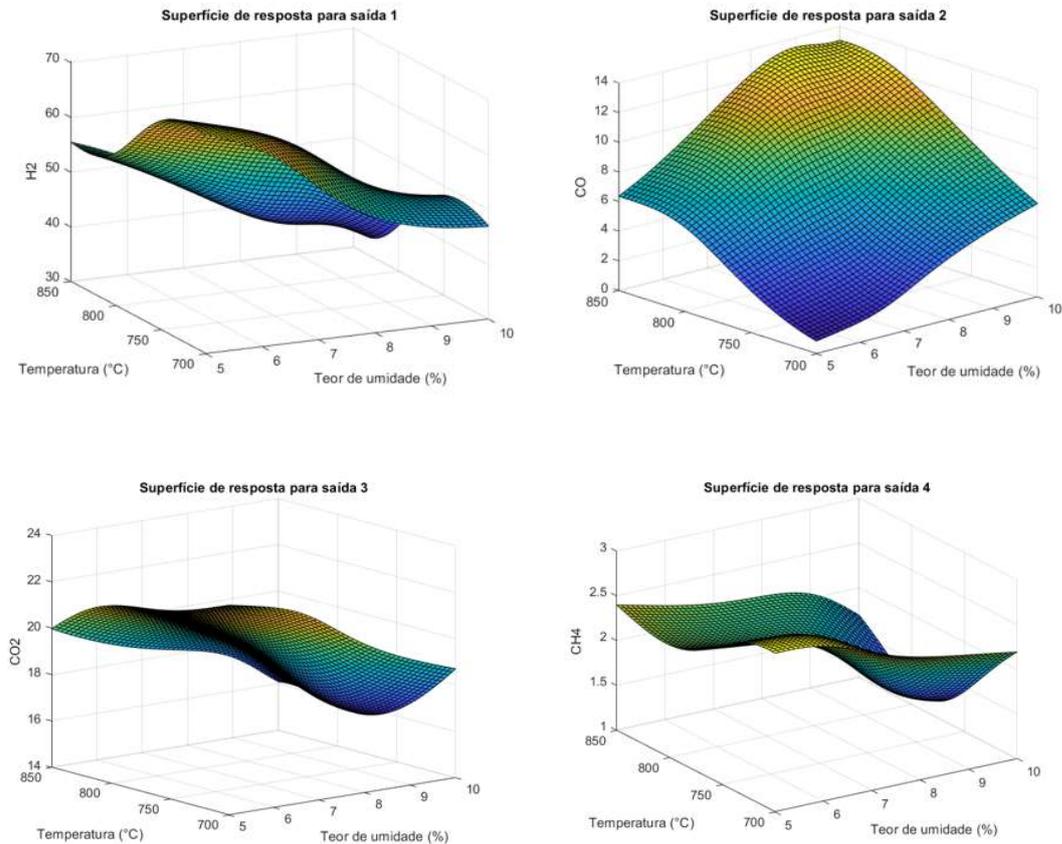


Figura 48 - Superfícies de resposta para as quatro saídas, com base no subgrupo de validação próximo à região bem treinada

De acordo com Sidek *et al.* (2020), o agente de gaseificação vapor é favorável para a produção de *syngas* rico em hidrogênio. Ademais, de acordo com Shahbaz *et al.* (2019), o vapor é o agente de gaseificação preferido para produção de *syngas* rico em hidrogênio. Comparando a imagem da saída de H<sub>2</sub> da Figura 22 com a imagem da saída de H<sub>2</sub> da Figura 48, observa-se que, na primeira, o valor de H<sub>2</sub> é de até cerca de 27,5 %, enquanto na segunda o H<sub>2</sub> vai até cerca de 70%. Isso corrobora com os ensinamentos de Sidek *et al.* (2020) e Shahbaz *et al.* (2019), uma vez que as superfícies de resposta da Figura 24 são referentes a um cenário com agente de gaseificação misto, predominando o ar, enquanto as superfícies da Figura 48 são referentes a um cenário em que o agente é o vapor.

A partir da imagem da saída CO da Figura 48, observa-se que quanto maior a temperatura e maior o teor de umidade, maior a produção de CO e quanto mais baixos os valores de temperatura e do teor de umidade, menor a produção de CO.

A partir da imagem da saída  $\text{CO}_2$  da Figura 48, observa-se maior produção de  $\text{CO}_2$  para menores valores de teor de umidade.

A partir da imagem da saída  $\text{CH}_4$  da Figura 48, observa-se que quanto mais baixo o valor de temperatura e mais baixo o teor de umidade, mais elevada é a produção de  $\text{CH}_4$ . Nota-se também que em aproximadamente 9,8% de teor de umidade e  $780^\circ\text{C}$  de temperatura está a região de menor produção de  $\text{CH}_4$ .

Como mencionado anteriormente, o *syngas* atua como matéria-prima para sínteses de biocombustíveis e produtos químicos e para geração de energia. Como exemplos da utilização do *syngas* pode-se listar a síntese de amônia, a produção de bio-hidrogênio e a geração de bio-metanol (Sikarwar et al., 2017). Nesse sentido, tendo em vista as superfícies de respostas apresentadas acima na Figura 48, as quais consideram vapor como agente de gaseificação, são aqui tecidos alguns comentários analíticos:

- (i) Caso o interesse de aplicação da biomassa seja na produção de bio-metanol, é interessante a obtenção de  $\text{H}_2$ ,  $\text{CO}$  e  $\text{CO}_2$ , tendo em vista as principais reações de equilíbrio de síntese do bio-metanol (vide equações (1) e (2) do presente trabalho). Nesse contexto, é recomendável ter, por exemplo, uma temperatura de gaseificação de cerca de  $760^\circ\text{C}$  e teor de umidade de cerca de 7%, posto que nessas condições há boas produções de  $\text{H}_2$ ,  $\text{CO}$  e  $\text{CO}_2$ , como pode ser visto na Figura 48;
- (ii) Caso o interesse de aplicação da biomassa seja na produção de combustível biometano, deve-se trabalhar com valores de temperatura de gaseificação próxima de  $700^\circ\text{C}$  e teor de umidade da biomassa próximo de 5%, como pode ser visto na Figura 48.

## 5. Conclusão

Para o desenvolvimento de modelos preditivos para a gaseificação de biomassa usando redes neurais artificiais, foi primeiramente necessária a construção de bancos de dados robustos compreendendo uma diversidade de biomassa, com diversos tipos e concentrações de agentes de gaseificação em diferentes tipos de gaseificadores. Para composição do banco de dados, foi realizada uma coleta de informações a partir de estudos anteriores disponíveis na literatura e tratamento dos dados obtidos. Após tratamento, resultou-se em 603 dados totais, os quais foram distribuídos em 491 dados no banco de dados principal, para treinamento e teste das RNA (cerca de 81,43 % dos dados totais) e 112 dados no banco de dados de validação (cerca de 18,57% dos dados totais).

O banco de dados principal foi dividido em dados de treinamento (80% - 393 dados) e dados de teste (20% - 98 dados). Com utilização do Matlab, 33 topologias de RNA foram testadas e, ao final, uma delas foi eleita como tendo a melhor topologia. A rede escolhida possuía 10 neurônios na camada de entrada; 2 camadas intermediárias, com funções de ativação logsig e 10 neurônios em cada camada intermediária; função de ativação purelin na camada final; 4 neurônios na camada final; e algoritmo de treinamento trainbr. Tal rede apresentou um bom desempenho, com valores de  $R^2$  de treinamento e de teste, respectivamente, maiores que 0,88 e 0,70, para cada uma das quatro saídas. Ademais, os valores de SSE de treinamento variaram de 0,2880 a 7,63 para as quatro saídas, sendo o menor de  $H_2$  e o maior de  $CH_4$ .

O segundo banco de dados foi utilizado na validação da RNA eleita como tendo a melhor topologia. O resultado da validação não foi muito satisfatório, uma vez que o  $R^2$  foi de 0,0025 a 0,1994 e o SSE de 2,9994 a  $2,42 \times 10^5$  para as quatro saídas, de modo que foi necessária uma reavaliação do comportamento dos dados de validação. No entanto, é fundamental destacar que estamos tratando, no presente trabalho, de banco de dados amplos, os quais englobam diferentes tipos de biomassa, de agente de gaseificação e de gaseificadores e essa diversidade de dados também é um dificultador para a robustez do modelo.

Nesse sentido, foram realizadas análises qualitativas e quantitativas dos dados de validação. Por meio da análise qualitativa, foi possível concluir que havia um

desbalanceamento na distribuição dos dados de validação das variáveis de entrada em relação aos dados de treinamento, principalmente C, H, O, cinzas e S/B, o que pode ser um dos motivos para o desempenho não satisfatório obtido na validação do modelo da RNA. Por meio da análise quantitativa, foi possível concluir que há um número considerável de pontos cujas distâncias entre dados de treinamento e de validação das variáveis de entrada são maiores do que o valor assumido como aceitável (0,6448), o que provavelmente também contribuiu para o desempenho ruim da validação.

Foi feita então uma avaliação da robustez na região de validação, por meio da qual pôde-se confirmar que seria possível selecionar um subgrupo de condições experimentais bem representativo dos dados de treinamento, usando o critério definido das distâncias, ou seja, menores distâncias entre dados de validação e região bem treinada. Assim, foram geradas superfícies de resposta, a partir da seleção de um subgrupo dos dados de treinamento bem representativos.

Conclui-se que o presente estudo alcançou seu objetivo de desenvolver um modelo de RNA para simular o processo de gaseificação de biomassa, capaz de prever a composição do gás produzido no referido processo com boa precisão. No entanto, houve uma falha na validação do modelo, proveniente de uma falha no banco de dados de validação, tendo em vista os dados de treinamento.

Como sugestão para trabalhos futuros, em relação a possíveis melhorias no modelo, podem ser citadas a reformulação do banco de dados de treinamento no sentido de incluir mais dados em regiões que atualmente existem classes vazias, bem como a reformulação do banco de dados de validação no sentido de incluir mais dados mais próximos de regiões bem treinadas. Portanto, como perspectivas futuras, a sugestão é aplicar técnicas de balanceamento de dados antes do fechamento dos bancos de dados e desenvolvimento dos modelos de RNA, por exemplo, reestruturando os dados e/ou coletando mais dados. A reestruturação dos dados pode ser realizada por amostragem dos dados de treinamento, teste e validação, utilizando análise estatística. Uma sugestão seria usar a ferramenta de estratificação de dados. Outras alternativas seriam considerar mais variáveis no modelo (por exemplo, tempo de retenção, material de leito, LHV do gás, diâmetro), de modo a torná-lo mais sensível, e/ou comparar o modelo de RNA desenvolvido com outras técnicas de modelagem (por exemplo, modelos considerando o equilíbrio termodinâmico).

## 6. Referências Bibliográficas

AGGARWAL, C. C. *Neural Networks and Deep Learning, A Textbook*, Springer, 2018.

AGGARWAL, K.K.; SINGH, Y.; CHANDRA, P.; PURI, M. Bayesian Regularization in a Neural Network Model to Estimate Lines of Code Using Function Points, **Journal of Computer Sciences**, v.1 (4), p. 505-509, 2005.

BALANÇO ENERGÉTICO NACIONAL, Relatório Síntese 2021 (Ano Base 2020), Empresa de Pesquisa Energética (EPE), 2021.

BARUAH, D.; BARUAH, D.C.; HAZARIKA, M. K. Artificial neural network based modeling of biomass gasification in fixed bed downdraft gasifiers. **Biomass and Bioenergy**, v. 98, p. 264-271, 2017.

BARUAH, D.; BARUAH, D.C. Modeling of biomass gasification: A review. **Renewable and Sustainable Energy Reviews**, v. 39, p. 806-815, 2014.

BASSYOUNI, M.; UL HASAN, S. W.; ABDEL-AZIZ, M. H.; ABDEL-HAMID, S.M.-S.; NAVEED, S.; HUSSAIN, A.; ANI, F. N. Date palm waste gasification in downdraft gasifier and simulation using ASPEN HYSYS. *Energy Conversion and Management*, v. 88, p. 693-699, 2014.

BURRA, K. R. G.; GUPTA, A. K. Modeling of biomass pyrolysis kinetics using sequential multi-step reaction model. **Fuel**, v. 237, p. 1057-1067, 2019.

CASTRO, V. E. M. Avaliação de estratégias para a gaseificação de resíduos de biomassa conforme aspectos logísticos e tecnológicos. **Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Planejamento Energético, COPPE, da Universidade Federal do Rio de Janeiro**. Março de 2020.

CHANG, A. C. C.; CHANG, H.; LIN, F.; LIN, K.; CHEN, C. Biomass gasification for hydrogen production. **International Journal of Hydrogen Energy**, v. 36, p. 14252-14260, 2011.

CHRISTY, J. V.; MOURAD, A. H. I. Biomass Gasification of Pongamia char in a Fluidized Bed Reactor. **6th International Conference on Renewable Energy: Generation and Applications (ICREGA)**, 2021.

DEEP LEARNING BOOK. **Capítulo 8 – Função de Ativação**. 2021. Disponível em <<https://www.deeplearningbook.com.br/funcao-de-ativacao/>>. Acesso em maio 2022.

GREEN, D. W.; PERRY, R. H. Perry's Chemical Engineers' Handbook. McGraw-Hill Education: Eighth Edition. 2400 pages, 2007.

EBADI, A. G.; HISORIEV, H. Gasification of algal biomass (*Cladophora glomerata* L.) with CO<sub>2</sub>/H<sub>2</sub>O/O<sub>2</sub> in a circulating fluidized bed. **Environmental Technology**, p. 749-755, 2017.

EICHLER, P.; SANTOS, F.; TOLEDO, M.; ZERBIN, P.; SCHMITZ, G.; ALVES, C.; RIES, L.; GOMES, F.; Produção do biometanol via gaseificação de biomassa lignocelulósica. **Química Nova**, v. 38, p. 828-835, 2015.

FAN, X.; YANG, L.; JIANG, J. Experimental study on industrial-scale CFB biomass gasification. **Renewable Energy**, v. 158, p. 32-36, 2020.

GAO, X.; ZHANG, Y.; BAO, F.; LI, B.; ZHAO, Y.; KE, C.; JIANG, B. CFD modeling of sawdust gasification in a lab-scale entrained flow reactor based on char intrinsic kinetics. Part 1: Model development. **Chemical Engineering & Processing: Process Intensification**, v. 125, p. 280-289, 2018.

GEORGE, J.; ARUN, P.; MURALEEDHARAN, C. Assessment of producer gas composition in air gasification of biomass using artificial neural network model. **International Journal of Hydrogen Energy**, v. 43, p. 9558-9568, 2018.

GUO, B.; LI, D.; CHENG, C.; LU, Z.; SHEN, Y. Simulation of biomass gasification with a hybrid neural network model. **Bioresource Technology**, v. 76, p. 77-83, 2001.

HABIBOLLAHZADE, A.; AHMADI, P.; ROSEN, M. A.; Biomass gasification using various gasification agents: Optimum feedstock selection, detailed numerical analyses and tri-objective grey wolf optimization. **Journal of Cleaner Production**, v. 284, p. 124718, 2021.

HAN, L.; WANG, Q.; YANG, Y.; YU, C.; FANG, M.; LUO, Z. Hydrogen production via CaO sorption enhanced anaerobic gasification of sawdust in a bubbling fluidized bed. **International Journal of Hydrogen Energy**, v. 36, p. 4820-4829, 2011.

JAHROMI, M. Y.; ATASHKARI, K.; KALTEH, M.; Comparison of different woody biomass gasification behavior in an entrained flow gasifier.

**Biomass Conversion and Biorefinery**, 2021.

JARUNGTHAMMACHOTE, S.; DUTTA, A.; Thermodynamic equilibrium model and second law analysis of a downdraft waste gasifier.

**Energy**, v. 32, p. 1660-1669, 2007.

KAMBLE, L.V.; PANGAVHANE, D. R.; SINGH, T. P.; Neural network optimization by comparing the performances of the training functions - Prediction of heat transfer from horizontal tube immersed in gas–solid fluidized bed. **International Journal of Heat and Mass Transfer**, v. 83, p. 337-344, 2015.

KHAN, A. A.; JONG, W.; JANSSENS, P. J.; SPLIETHOFF, H. Biomass combustion in fluidized bed boilers: Potential problems and remedies. **Fuel Processing Technology**, v. 90, p. 21-50, 2009.

KIBRIA M. A.; SRIPADA, P.; WOO, M. W.; BHATTACHARYA, S. Fate of a biomass particle during CO<sub>2</sub> gasification: a mathematical model under entrained flow condition at high temperature. **Energy**, v. 168, p. 1045-1062, 2019.

KIRUBAKARAN, V.; SIVARAMAKRISHNAN, V.; NALINI, R.; SEKAR, T.; PREMALATHA, M.; SUBRAMANIAN, P. A review on gasification of biomass. **Renewable and Sustainable Energy Reviews**, v. 13, p. 179-186, 2009.

KOPPATZ, S.; PFEIFER, C.; RAUCH, R.; HOFBAUER, H.; MARQUARD-MOELLENSTEDT, T.; SPECHT, M. H<sub>2</sub> rich product gas by steam gasification of biomass with in situ CO<sub>2</sub> absorption in a dual fluidized bed system of 8 MW fuel input. **Fuel Processing Technology**, v. 90, p. 914-921, 2009.

LA VILLETTE, M.; COSTA, M.; MASSAROTTI, N. Modelling approaches to biomass gasification: A review with emphasis on the stoichiometric method. **Renewable and Sustainable Energy Reviews**, v. 74, p. 71-88, 2017.

LI, J.; YIN, Y.; ZHANG, X.; LIU, J.; YAN, R.; Hydrogen-rich gas production by steam gasification of palm oil wastes over supported tri-

metallic catalyst. **International Journal of Hydrogen Energy**, v. 34, p. 9108-9115, 2009.

LIU, W.; TIAN, Y.; YAN, H.; ZHOU, X.; TAN, Y.; YANG, Y.; LI, Z.; YUAN, L.; Gasification of biomass using oxygen-enriched air as gasification agent: a simulation study. **Biomass Conversion and Biorefinery**, p. 1-8, 2021.

LOHA, C.; CHATTOPADHYAY, H.; CHATTERJEE, P. K. Energy generation from fluidized bed gasification of rice husk. **Journal of Renewable Sustainable Energy**, v. 5, p. 043111, 2013.

LOHA, C.; CHATTOPADHYAY, H.; CHATTERJEE, P. K. Thermodynamic analysis of hydrogen rich synthetic gas generation from fluidized bed gasification of rice husk. **Energy**, v. 36, p. 4063-4071, 2011.

LV, P.; YUAN, Z.; WU, C.; MA, L.; CHEN, Y.; TSUBAKI, N. Bio-syngas production from biomass catalytic gasification. **Energy Conversion and Management**, v. 48, p. 1132–1139, 2007.

MICHAELIDES, E. E. *Alternative Energy Sources*, Editora Springer, 2012.

NETO, J. G.; OZORIO, L. V.; ABREU, T. C. C.; SANTOS, B. F.; PRADELLE, F. Modeling of biogas production from food, fruits and vegetables wastes using artificial neural network (ANN), **Fuel**, v. 285, p. 119081, 2021.

NGO, S. I.; Nguyen, D. B.; Lim, Y.; Song, B.; Lee, U.; Choi, Y.; Song, J. Performance evaluation for dual circulating fluidized-bed steam gasifier of biomass using quasi-equilibrium three-stage gasification model. **Applied Energy**, v. 88, p. 5208–5220, 2011.

OZONOH, M.; OBOIRIEN, B. O.; HIGGINSON, A.; DARAMOLA, M. O. Performance evaluation of gasification system efficiency using artificial neural network. **Renewable Energy**, v. 145, p. 2253-2270, 2020.

PANDEY, D. S.; DAS, S.; PAN, I.; LEAHY, J. J.; KWAPINSKI, W. Artificial neural network based modelling approach for municipal solid waste gasification in a fluidized bed reactor. **Waste Management**, v. 58, p. 202-213, 2016.

PFEIFER, C.; PUCHNER, B.; HOFBAUER, H. Comparison of dual fluidized bed steam gasification of biomass with and without selective

transport of CO<sub>2</sub>. **Chemical Engineering Science**, v. 64, p. 5073-5083, 2009.

PLIS, P.; WILK, R. K. Theoretical and experimental investigation of biomass gasification process in a fixed bed gasifier. **Energy**, v. 36, p. 3838-3845, 2011.

PUIG-ARNAVAT, M.; BRUNO, J. C.; CORONAS, A. Review and analysis of biomass gasification models. **Renewable and Sustainable Energy Reviews**, v. 14, p. 2841-2851, 2010.

PUIG-ARNAVAT, M.; HERNÁNDEZ, J. A.; BRUNO, J. C.; CORONAS, A. Artificial neural network models for biomass gasification in fluidized bed gasifiers. **Biomass and Bioenergy**, v. 49, p. 279-289, 2013.

RAPAGNÀ, S.; JAND, N.; KIENNEMANN, A.; FOSCOLO, P. U. Steam-gasification of biomass in a fluidized-bed of olivine particles. **Biomass and Bioenergy**, v. 19, p. 187-197, 2000.

SAMADI, S. H.; GHOBADIAN, B.; NOSRATI, M. Prediction and estimation of biomass energy from agricultural residues using air gasification technology in Iran. **Renewable Energy**, v. 149, p. 1077-1091, 2020.

SÁNCHEZ, C. G. Tecnologia da Gaseificação de Biomassa, Editora Átomo, 2010.

SANSANIWAL, S. K.; PAL, K.; ROSEN, M. A.; TYAGI, S. K. Recent advances in the development of biomass gasification technology: A comprehensive review. **Renewable and Sustainable Energy Reviews**, v. 72, p. 363-384, 2017.

SANSANIWAL, S. K.; ROSEN, M. A.; TYAGI, S. K. Global challenges in the sustainable development of biomass gasification: An overview. **Renewable and Sustainable Energy Reviews**, v. 80, p. 23-43, 2017.

SERRANO, D.; GOLPOUR, I.; SÁNCHEZ-DELGADO, S. Predicting the effect of bed materials in bubbling fluidized bed gasification using artificial neural networks (ANNs) modeling approach. **Fuel**, v., 266, p. 117021, 2020.

SHAHBAZ, M.; TAQVI, S. A.; LOY, A. C.M.; INAYAT, A.; UDDIN, F.; BOKHARI, F.; NAQVI, S. R. Artificial neural network approach for the steam

gasification of palm oil waste using bottom ash and CaO, **Renewable Energy**, v. 132, p. 243-254, 2019.

SHEELA, K. G.; DEEPA, S. N. Review on Methods to Fix Number of Hidden Neurons in Neural Networks. **Hindawi Publishing Corporation**, v. 2013, p. 1-11, 2013.

SIDEK, F. N.; SAMAD, N. A. F.; SALEH, S.; Review on effects of gasifying agents, temperature and equivalence ratio in biomass gasification process. **Materials Science and Engineering**, 863, 2020.

SIKARWAR, V. S.; ZHAO, M.; FENNELL, P.; SHAH, N.; ANTHONY, E. J. Progress in biofuel production from gasification, **Progress in Energy and Combustion Science**, v. 61, p. 189-248, 2017.

SIMONE, M.; BARONTINI, F.; NICOLELLA, C.; TOGNOTTI, L. Gasification of pelletized biomass in a pilot scale downdraft gasifier. **Bioresource Technology**, v. 116, p. 403-412, 2012.

SREEJITH, C. C.; MURALEEDHARAN, C.; ARUN, P. Performance prediction of fluidised bed gasification of biomass using experimental data-based simulation models. **Biomass Conversion and Biorefinery**, v. 3, p. 283-304, 2013.

SUSASTRIAWAN, A. A. P.; SAPTOADIB, H.; PURNOMOB. Small-scale downdraft gasifiers for biomass gasification: A review. **Renewable and Sustainable Energy Reviews**, v. 76, p. 989-1003, 2017.

TRNINIĆ, M.; STOJILJKOVIĆ, D.; MANIĆ, N.; SKREIBERG, O.; WANG, L.; JOVOVIĆ, A.; A mathematical model of biomass downdraft gasification with an integrated pyrolysis model. **Fuel**, v. 265, p. 116867, 2020.

VASSILEV, S. V.; BAXTER, D.; ANDERSEN, L. K.; VASSILEVA, C. G. An overview of the chemical composition of biomass. **Fuel**, v. 89, p. 913-933, 2010.

WANG, H.; CHAFFART, D.; RICARDEZ-SANDOVAL, L. A.; Modelling and optimization of a pilot-scale entrained-flow gasifier using artificial neural networks. **Energy**, v. 188, p. 116076, 2019.

WEERACHANCHAI, P.; HORIO, M.; TANGSATHITKULCHAI, C. Effects of gasifying conditions and bed materials on fluidized bed steam

gasification of wood biomass. **Bioresource Technology**, v. 100, p. 1419-1427, 2009.

WOJNICKA, B.; ŚCIAŻKO, M.; SCHMID, J. C.; Modelling of biomass gasification with steam. **Biomass Conversion and Biorefinery**, v. 11, p. 1787-1805, 2021.

XIAO, G.; NI, M.; CHI, Y.; JIN, B.; XIAO, R.; ZHONG, Z.; HUANG, Y. Gasification characteristics of MSW and an ANN prediction model, **Waste Management**, v. 29, p. 240-244, 2009.

XU, G.; MURAKAMI, T.; SUDA, T.; MATSUZAWA, Y.; TANI, H. Two-stage dual fluidized bed gasification: Its conception and application to biomass. **Fuel Processing Technology**, v. 90, p. 137-144, 2009.

ZAINAL, Z. A.; ALI, R.; LEAN, C. H.; SEETHARAMU, K. N.; Prediction of performance of a downdraft gasifier using equilibrium modeling for different biomass material. **Energy Conversion & Management**, v. 42, p. 1499-1515, 2001.

ZHANG, B.; ZHANG, L.; YANG, Z.; YAN, Y.; PU, G.; GUO, M.; Hydrogen-rich gas production from wet biomass steam gasification with CaO/MgO. **International Journal of Hydrogen Energy**, v. 40, p. 8816-8823, 2015.

ZHOU, J.; CHEN, Q.; ZHAO, H.; CAO, X.; MEI, Q.; LUO, Z.; CEN, K. Biomass–oxygen gasification in a high-temperature entrained-flow gasifier. **Biotechnology Advances**, v. 27, p. 606-611, 2009.

## Anexo I

Tabela 8 – Descrição de diversos algoritmos de treinamento (adaptado de Aggarwal, 2005)

<b>Algoritmo de treinamento</b>	<b>Descrição</b>
trainb	Treina uma rede com regras de aprendizado de peso e bias com atualizações em lote. Os pesos e bias são atualizados no final de uma passagem inteira pelos dados de entrada.
trainbfg	Atualiza os valores de peso e bias de acordo com o método BFGS quasi-Newton.
trainbr	Atualiza os valores de peso e bias de acordo com a otimização de Levenberg-Marquardt. Ele minimiza uma combinação de erros quadráticos e pesos e então determina a combinação correta para produzir uma rede que generaliza adequadamente. O processo é chamado de regularização Bayesiana ( <i>Bayesian Regularisation</i> – BR).
trainc	Treina uma rede com regras de aprendizado de peso e bias com atualizações incrementais após cada apresentação de uma entrada. As entradas são apresentadas em ordem cíclica.
train cgb	Atualiza os valores de peso e bias de acordo com a retropropagação do gradiente conjugado com reinicializações de Powell-Beale.
train cgf	Atualiza os valores de peso e bias de acordo com a retropropagação do gradiente conjugado com atualizações de Fletcher-Reeves.
train cgp	Atualiza os valores de peso e bias de acordo com a retropropagação de gradiente conjugado com atualizações de Polak-Ribiere.
train gd	Atualiza os valores de peso e bias de acordo com a descida do gradiente.
train gda	Atualiza os valores de peso e bias de acordo com a descida do gradiente com taxa de aprendizado adaptável.
train gdm	Atualiza os valores de peso e bias de acordo com a descida do gradiente com momento.
train gdx	Atualiza os valores de peso e bias de acordo com o momento de descida do gradiente e uma taxa de aprendizado adaptável.
train lm	Atualiza os valores de peso e bias de acordo com a otimização de Levenberg-Marquardt.
train oss	Atualiza os valores de peso e bias de acordo com o método secante de uma etapa.
train rp	Atualiza os valores de peso e bias de acordo com o algoritmo de retropropagação resiliente (RPROP).
train scg	Atualiza os valores de peso e bias de acordo com o método de gradiente conjugado escalonado.

## Apêndice I - Trabalhos que utilizam RNA para simulação de gaseificação de biomassa

Referência	Agente de gaseificação	Tipo de gaseificador	Biomassa	Variáveis de entrada	Variáveis de saída	Descrição da RNA									
						Nº de casos	% Treino	% Teste	% Validação	Algoritmo de treinamento	Função de ativação	Nº de neurônios na camada de entrada	Nº de camadas intermediárias	Nº de neurônios nas camadas intermediárias	Nº de neurônios na camada de saída
Ozonoh, 2019	Ar	Vários (por ex, leitos arrastado, fluidizado e fixo)	Não especificado	C, H, O, N, S, Cinzas, VM, MC, LHV, ER, Temperatura	CGE(%), CCE(%), LHV <sub>gás</sub> , Rendimento de gás, CO, CO <sub>2</sub> , CH <sub>4</sub> , H <sub>2</sub>	315	70	15	15	trainbr e trainlm	tansig	11	1	10, 20, 21, 22 e 23	8
George, 2018	Ar	Leito fluidizado borbulhante	Serragem, casca de coco, casca de café, bagaço de cana e casca de amendoim	C, H, O, Cinzas, MC, ER, Temperatura	H <sub>2</sub> , CO, CH <sub>4</sub> , CO <sub>2</sub>	70	70	15	15	trainlm	tansig e purelin	7	1	15	4

Referência	Agente de gaseificação	Tipo de gaseificador	Biomassa	Variáveis de entrada	Variáveis de saída	Descrição da RNA									
						Nº de casos	% Treino	% Teste	% Validação	Algoritmo de treinamento	Função de ativação	Nº de neurônios na camada de entrada	Nº de camadas intermediárias	Nº de neurônios nas camadas intermediárias	Nº de neurônios na camada de saída
Pandey, 2016	Ar	Leito fluidizado	Resíduos sólidos municipais	C, H, O, N, S, Cinzas, MC, ER, Temperatura	Rendimento de gás, $LHV_{gás}$ e $LHV_P$ (dos produtos de gaseif)	67	70	15	15	trainlm	tansig, logsig e purelin	9	1 e 2	30 (camada única) e 15 (camada dupla)	3
Serrano, 2020	Ar	Leito fluidizado borbulhante	Não especificado	C, H, O, Cinzas, MC, ER, S/B, material de leito, Temperatura	$CO_2$ , $CO$ , $CH_4$ , $H_2$ e rendimento de gás	203	80 (162 casos)	Outros 10 dados usados para teste	20 (41 casos)	trainlm	purelin, tansig e logsig	9	1 e 2	1 a 30	1 (uma RNA para cada uma das 5 saídas)
Sreejith, 2013	Ar-vapor	Leito fluidizado	Serragem de madeira	ER, S/B e Temperatura	Temperatura, $CO_2$ , $CO$ , $H_2$ , $CH_4$ , $N_2$	-	90	10	-	trainlm	tansig	3 e 2	1	12 e 24	1 e 5
Puig-Arnavat, 2013	Ar e ar-vapor	Leito fluidizado borbulhante (BFB) e leito fluidizado	Madeira	C, H, O, Cinzas, MC, ER, S/B, Temperatura	Rendimento de gás, $CO$ , $CO_2$ , $H_2$ , $CH_4$	18 (CFB) e 36 (BFB)	80	20 (os grupos de validação e teste eram os mesmos)		trainlm	tansig e purelin	7 (CFB) e 8 (BFB)	1	2	1

Referência	Agente de gaseificação	Tipo de gaseificador	Biomassa	Variáveis de entrada	Variáveis de saída	Descrição da RNA									
						N° de casos	% Treino	% Teste	% Validação	Algoritmo de treinamento	Função de ativação	N° de neurônios na camada de entrada	N° de camadas intermediárias	N° de neurônios nas camadas intermediárias	N° de neurônios na camada de saída
		circulante (CFB)													
Baruah, 2017	-	Leito fixo co-corrente	Matéria-prima lenhosa	C, H, O, Cinzas, MC, Temperatura	CH <sub>4</sub> , CO, CO <sub>2</sub> e H <sub>2</sub>	63	70	30		trainlm	tansig e purelin	6	1	5 (CH <sub>4</sub> e CO), 4 (CO <sub>2</sub> ), 3 (H <sub>2</sub> )	1
Xiao, 2008	-	Leito fluidizado	Resíduos sólidos municipais	ER, percentual de biomassa, Temperatura	LHV <sub>gás</sub> , LHV (dos produtos de gaseif) e rendimento de gás	91	89	-	10,989 (10 resultados)	-	logsig, tansig e purelin	7	1	5	3
Muhammad Shahbaz, 2019	Vapor	Leito fluidizado e leito fixo	Resíduos de óleo de palma	S/B, razão Cao/biomassa, % em peso de cinzas de carvão, Temperatura	H <sub>2</sub> , CO <sub>2</sub> , CO, CH <sub>4</sub> , rendimento de gás, LHV <sub>gás</sub> e HHV <sub>gás</sub>	-	80	20	20	trainlm	-	2	1	7	1

Referência	Agente de gaseificação	Tipo de gaseificador	Biomassa	Variáveis de entrada	Variáveis de saída	Descrição da RNA									
						N° de casos	% Treino	% Teste	% Validação	Algoritmo de treinamento	Função de ativação	N° de neurônios na camada de entrada	N° de camadas intermediárias	N° de neurônios nas camadas intermediárias	N° de neurônios na camada de saída
Guo, 2001	Vapor	Leito fluidizado	Serragem de choupo, serragem de pinus, bagaço de cana triturado e talo de algodão	Tempo e Temperatura	CO, CO <sub>2</sub> , CH <sub>4</sub> , H <sub>2</sub>	-	-	-	-	-	logsig	2	1	5	4

## Apêndice II – Cálculo do LHV

De acordo com a tabela 2-179 de Perry:

Composto	Entalpia de combustão líquida padrão, J/kmol x 1E-09
H <sub>2</sub>	-0,24182
CO	-0,283
CO <sub>2</sub>	--
CH <sub>4</sub>	-0,80262

Primeiramente, foi assumido gás ideal, de forma a considerar que %vol = %mol.

Gás: 6% molar H<sub>2</sub> + 16,1% molar CO + 6,3% molar CH<sub>4</sub> (valores variáveis de acordo com o banco de dados)

Base: 1 mol de gás → 0,06 mol H<sub>2</sub> + 0,161 mol CO + 0,063 mol CH<sub>4</sub>

Tabelado:

[H<sub>2</sub>]: 0,24182 x 10<sup>6</sup> J/mol ... H<sub>2</sub> + ½ O<sub>2</sub> → H<sub>2</sub>O(g)

[CO]: 0,283 x 10<sup>6</sup> J/mol ... CO + ½ O<sub>2</sub> → CO<sub>2</sub>(g)

[CH<sub>4</sub>]: 0,80262 x 10<sup>6</sup> J/mol ... CH<sub>4</sub> + 2O<sub>2</sub> → CO<sub>2</sub> + 2 H<sub>2</sub>O

(0,06 x 0,24182 + 0,161 x 0,283 + 0,063 x 0,80262) x 10<sup>6</sup> = LHV [J/mol de gás]

LHV [J/mol de gás] / 22,4 = LHV [kJ/Nm<sup>3</sup>]

Fator 22,4 veio do seguinte raciocínio:

Nm<sup>3</sup>: Normal metro cúbico, condições normais (T = 273, 15 K e P = 101325 Pa)

PV = nRT → 101325 x V = 1 (mol de gás) x 8,314 x 273,15 → V = 22,4 x 10<sup>-3</sup> m<sup>3</sup>

LHV [J/mol de gás] x [1 mol de gás / (22,4 x 10<sup>-3</sup> m<sup>3</sup>)] x (10<sup>-3</sup> kJ / 1 J) = LHV [kJ/Nm<sup>3</sup>]

## Apêndice III – Cálculo do N<sub>2</sub> e demais gases

Para os dados em que os percentuais volumétricos dos gases de H<sub>2</sub>, CO, CO<sub>2</sub> e CH<sub>4</sub> eram livres de N<sub>2</sub>, fez-se o cálculo do N<sub>2</sub>, bem como dos demais gases considerando o nitrogênio, para complementar o banco de dados.

Para realização dos cálculos, utilizou-se os dados do próprio banco de dados. Como exemplo, considera-se o dado em que: C = 50 %wt db, H = 5,70 %wt db, O = 44,10 %wt db, N = 0, ER = 0,32, H<sub>2</sub> = 18,09 %vol, CO = 36,18 %vol, CO<sub>2</sub> = 34,88 %vol, e CH<sub>4</sub> = 7,75 %vol.

Base de cálculo: 100 g de biomassa

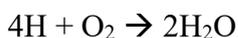
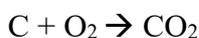
Pela porcentagem mássica (50% de C, 5,7% de H e 44,10% de O) e a base de 100 g de biomassa, obtemos:

50 g de C + 5,7 g de H e 44,10 g de O

Com as massas moleculares (12 g/mol para o carbono; 1 g/mol para o H e 16 g/mol para o O) é possível calcular os números de mols:

4,17 mols de C + 5,7 mols de H + 2,76 mols de O

Reações utilizadas:



Pelas reações:

1 mol de C – 1 mol de O<sub>2</sub>

4,17 mols de C – x mols de O<sub>2</sub>

x = 4,17 mols O<sub>2</sub>

4 mols de H – 1 mol de O<sub>2</sub>

5,7 mols de H – y mols de O<sub>2</sub>

y = 1,425 mol O<sub>2</sub>

TOTAL de O<sub>2</sub> utilizado nas reações = x + y = 4,17 + 1,425 = 5,595 mols de O<sub>2</sub>

É necessário descontar a quantidade de O da biomassa para calcular a quantidade de O<sub>2</sub> na estequiometria. Em seguida, a partir do O<sub>2</sub> estequiométrico, calcular a quantidade de O<sub>2</sub> real, multiplicando pelo ER, conforme a seguir.

TOTAL de O<sub>2</sub> ar = 5,595 – (2,76/2) = 4,215 mols de O<sub>2</sub> (estequiométrico)

ER = (O<sub>2</sub>)<sub>real</sub> / (O<sub>2</sub>)<sub>estequiométrico</sub>

$$(O_2)_{\text{real}} = ER \times (O_2)_{\text{estequiométrico}} = 0,32 \times 4,215 = 1,3488 \text{ mols de } O_2 \text{ (real)}$$

Considerando que o ar é composto por 21% de  $O_2$  e 79% de  $N_2$ , obtem-se a quantidade de AR:

$$1,3488 \text{ mols} / 0,21 = 6,4228 \text{ mols de AR}$$

Encontrando a quantidade de  $N_2$ :

$$6,4228 \text{ mols de AR} \times 0,79 = 5,074 \text{ mols de } N_2$$

Em seguida, foi feito um balanço de elementos:

BIOMASSA (C + H + O) desprezando N e S C: 4,17 mols C	+	UMIDADE (H + O)	+	AR (O + N) N: 10,148 mols N (5,074*2)	=	SYNGAS ( $H_2$ , CO, $CO_2$ , $CH_4$ – livres de $N_2$ ) C: 4,17 mols C ( $CO$ + $CO_2$ + $CH_4$ )
--	---	--------------------	---	---	---	--

$n_s$ : numero de mols do syngas livre de  $N_2$

$$C: 0,3618 * n_s (CO) + 0,3488 * n_s (CO_2) + 0,0775 * n_s = 4,17 \rightarrow n_s = 5,29$$

$$n_{N_2} = 5,074 \text{ mols (calculado anteriormente)}$$

$$n_T = n_s + n_{N_2} = 5,29 + 5,074 = 10,364$$

$$\%N_2 = (n_{N_2} / n_T) * 100 = (5,074 / 10,364) * 100 = 48,96 \%$$

No banco de dados, deve-se substituir os valores de  $H_2$ , CO,  $CO_2$  e  $CH_4$  livres de  $N_2$  por valores que consideram o  $N_2$ , para que todos os gases estejam na mesma base. Dessa forma, tem-se os seguintes cálculos:

$$n_{H_2} = 5,29 * 0,1809 = 0,957 \text{ mols}$$

$$n_{CO} = 5,29 * 0,3618 = 1,914 \text{ mols}$$

$$n_{CO_2} = 5,29 * 0,3488 = 1,845 \text{ mols}$$

$$n_{CH_4} = 5,29 * 0,0775 = 0,41 \text{ mols}$$

$$n_T = n_{H_2} + n_{CO} + n_{CO_2} + n_{CH_4} + n_{N_2} = 10,2$$

Percentuais, considerando que **não** são livres de niterogênio:

$$\%H_2 = (n_{H_2} / n_T) * 100 = 9,38 \%$$

$$\%CO = (n_{CO} / n_T) * 100 = 18,77 \%$$

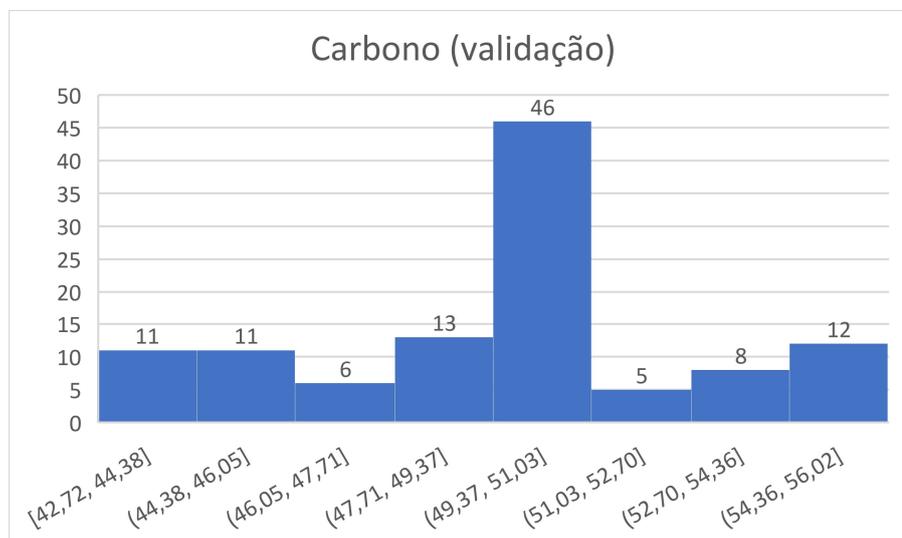
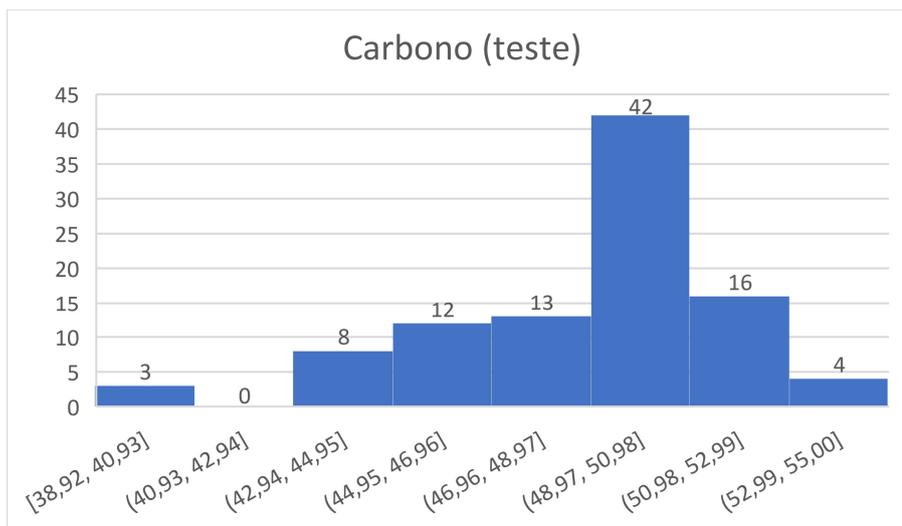
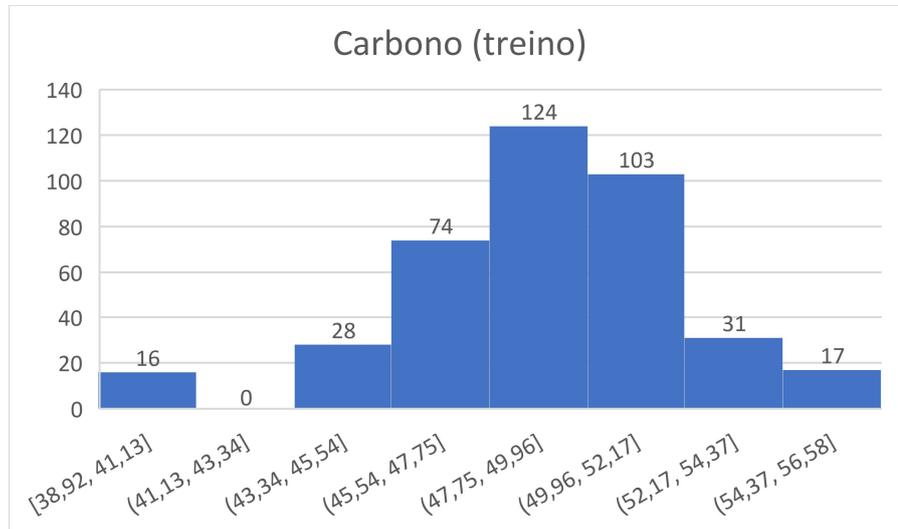
$$\%CO_2 = (n_{CO_2} / n_T) * 100 = 18,09 \%$$

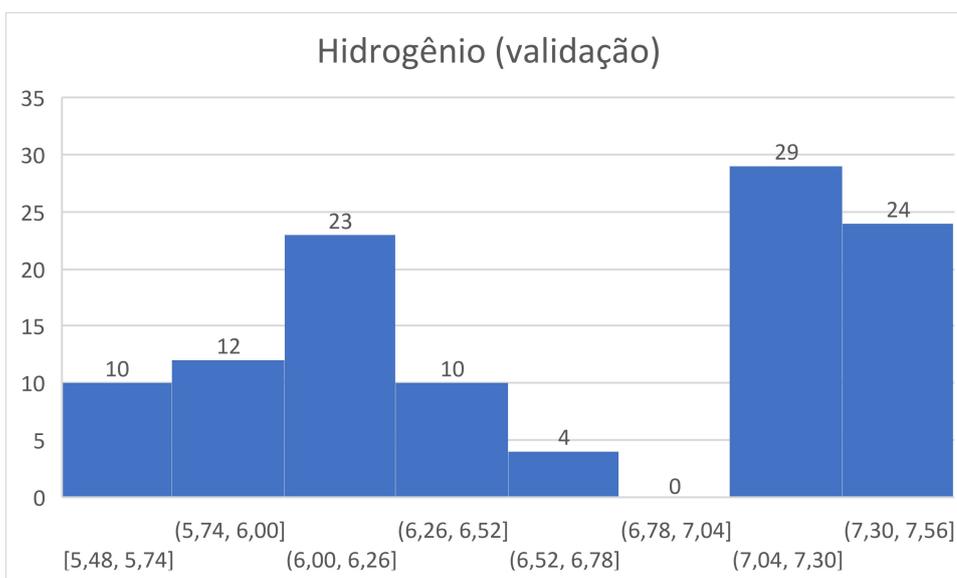
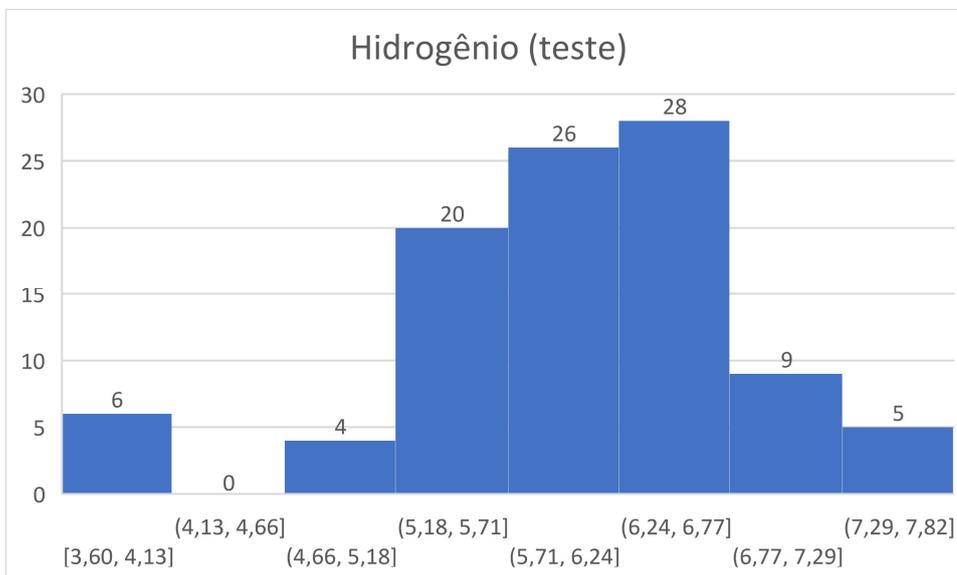
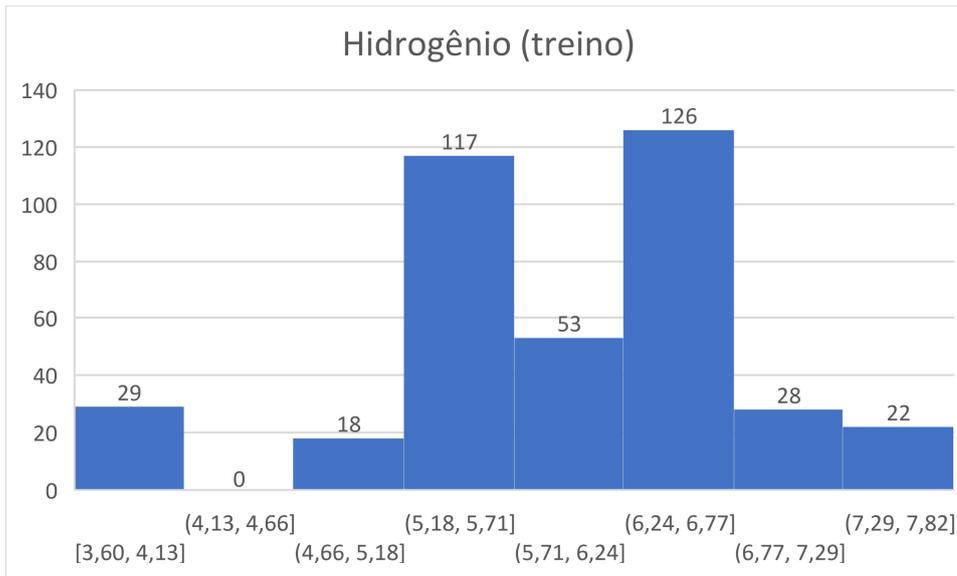
$$\%CH_4 = (n_{CH_4} / n_T) * 100 = 4,02 \%$$

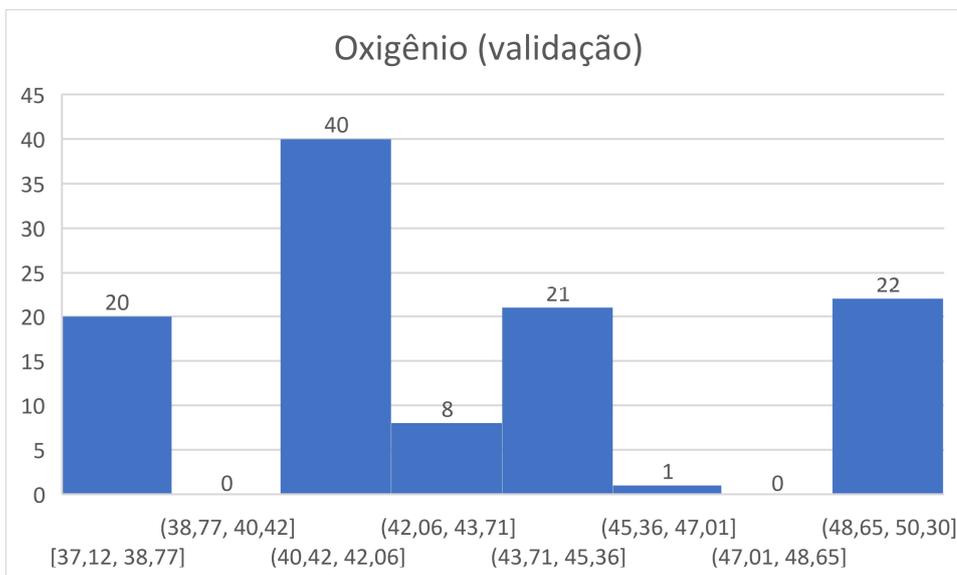
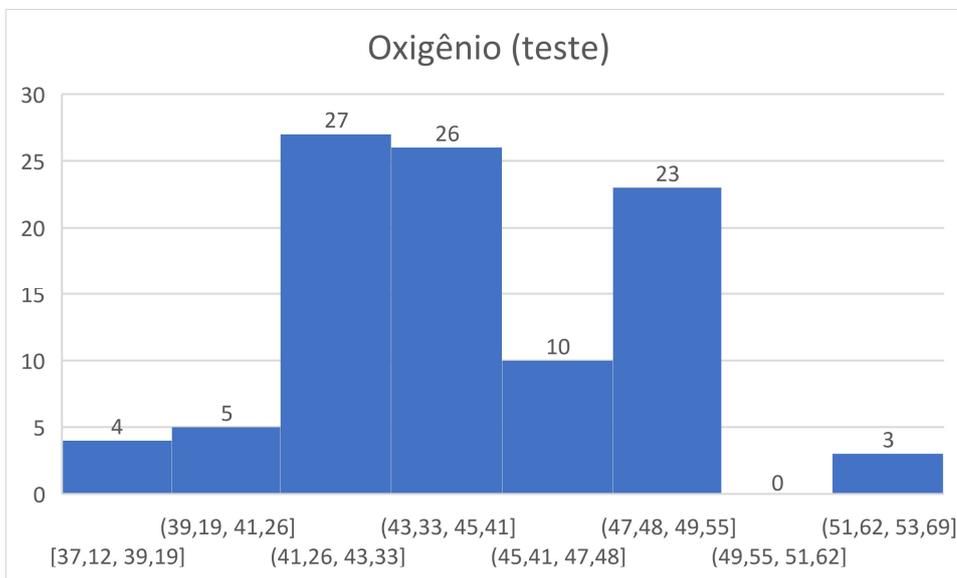
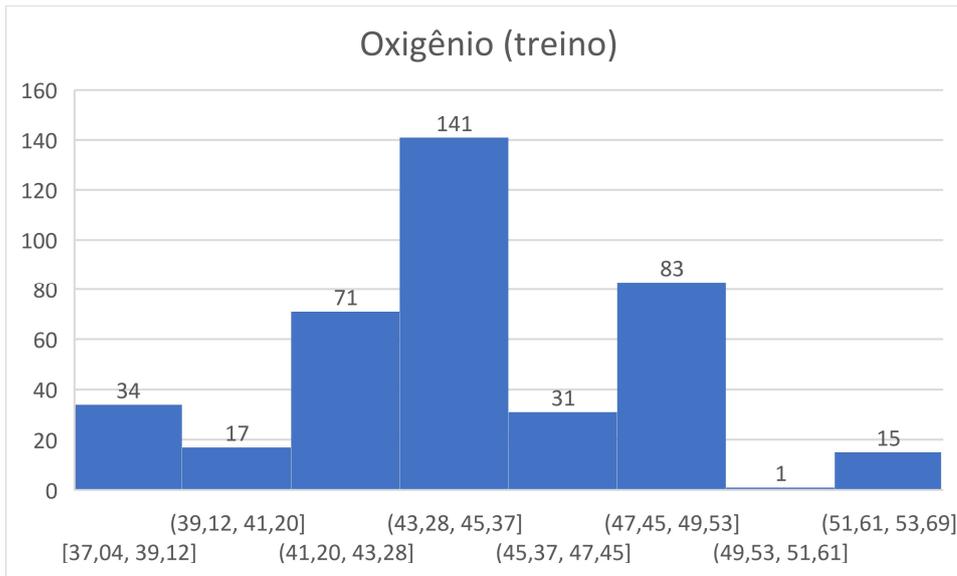
Acrescentando o %N<sub>2</sub> calculado acima, tem-se a composição em base seca, ou seja, livre de água, que foi utilizada no banco de dados. É uma base comumente empregada para fazer as contas em gaseificação.

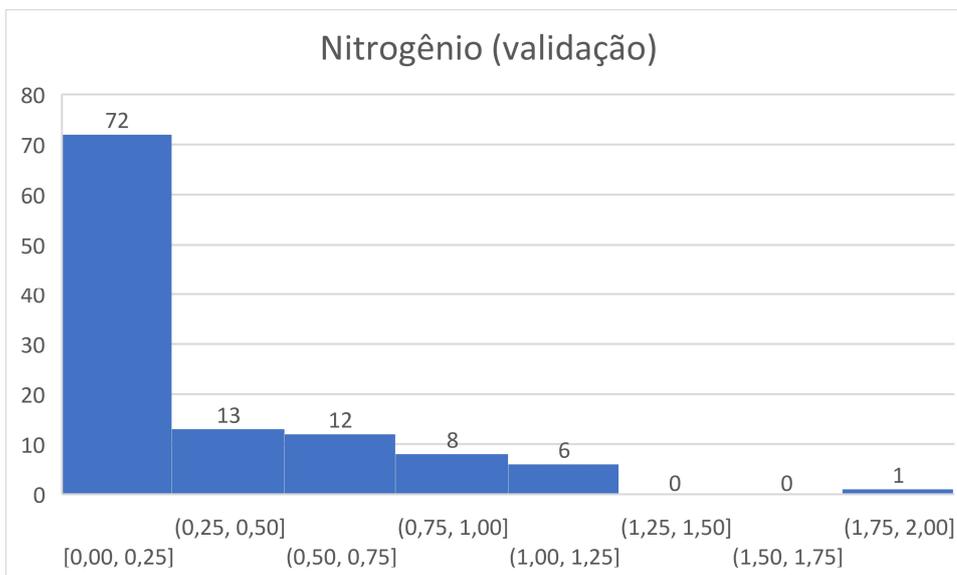
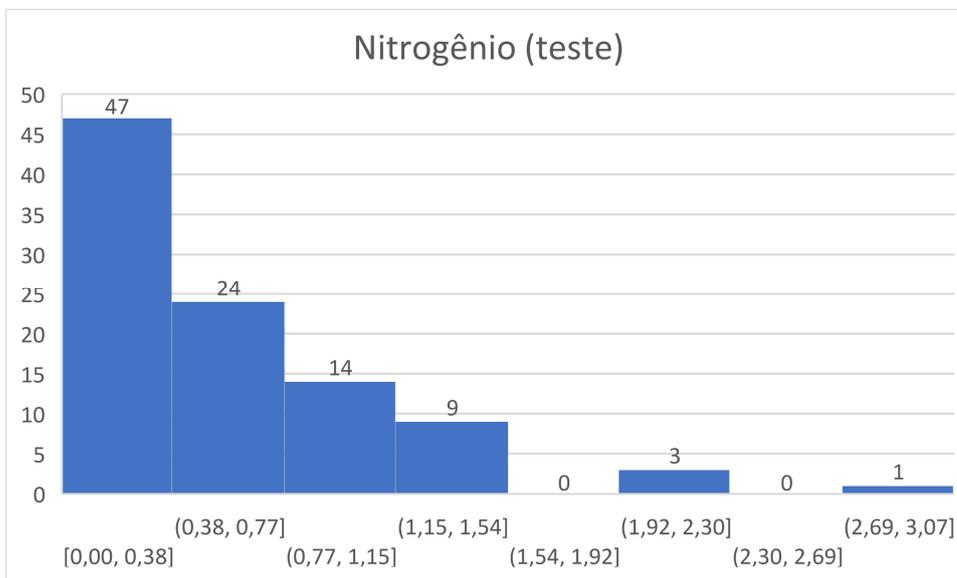
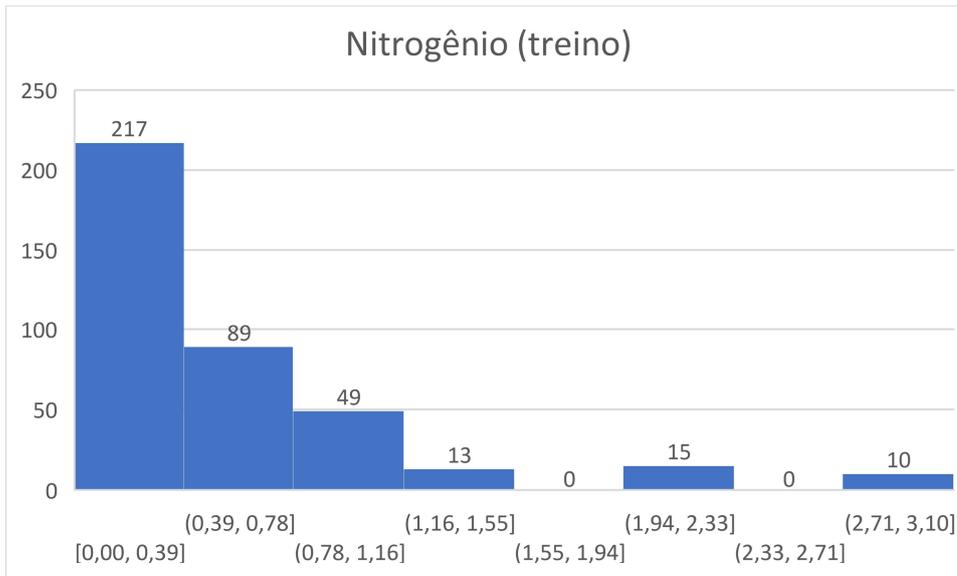
No exemplo aqui fornecido, a biomassa tinha um valor de N próximo de zero. Todavia, ressalta-se que o percentual em peso de N, quando não nulo, foi levado em consideração nos cálculos no banco de dados.

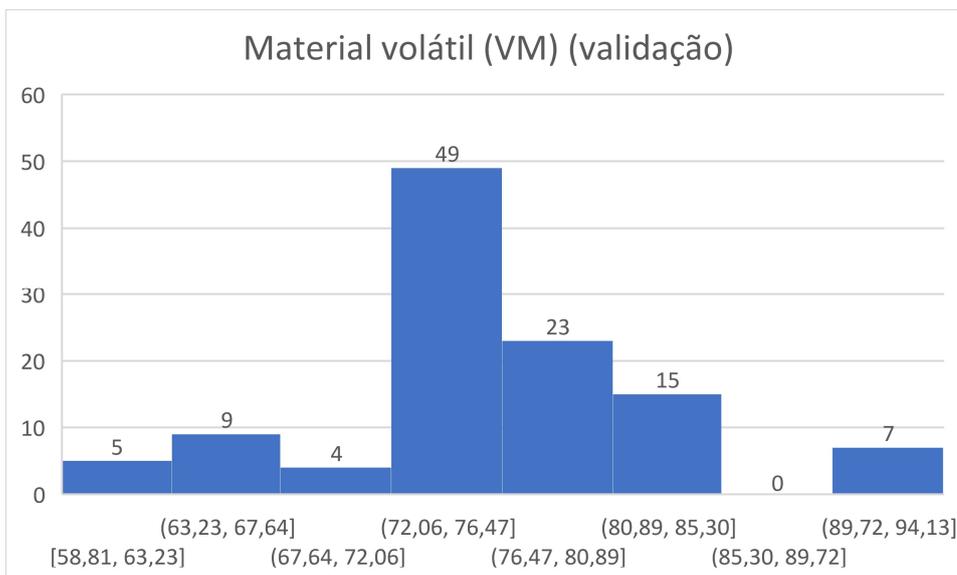
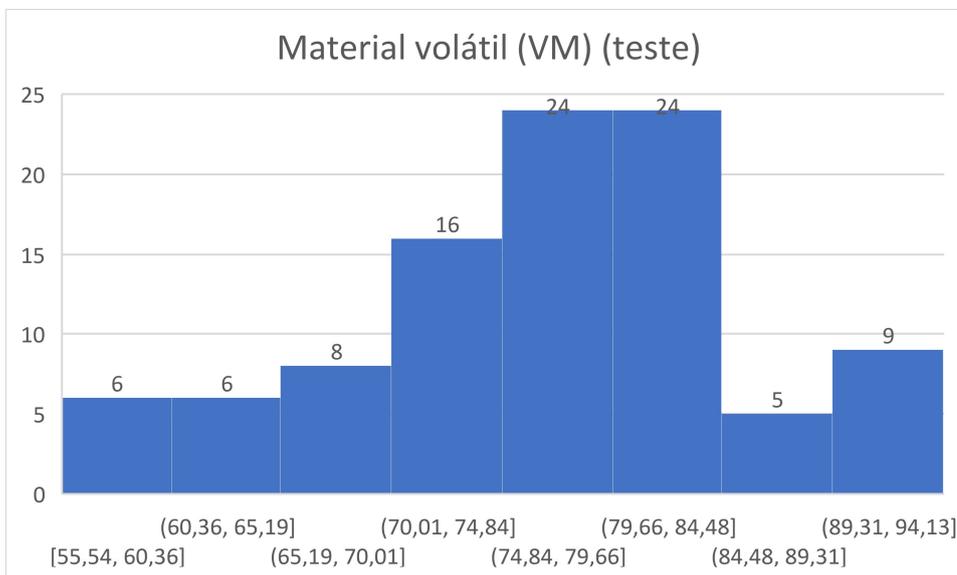
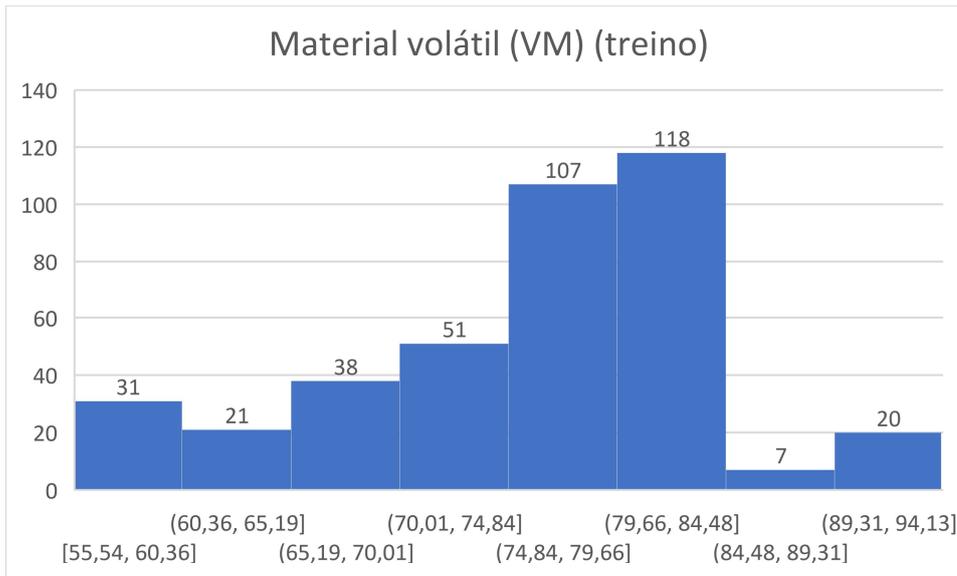
## Apêndice IV – Histogramas das variáveis de entrada dos dados para treinamento, teste e validação das RNA

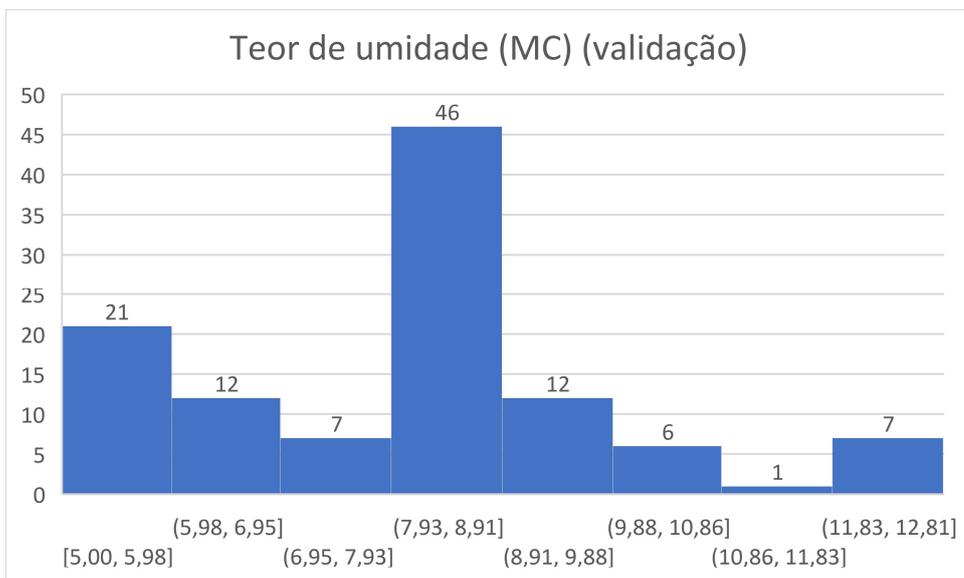
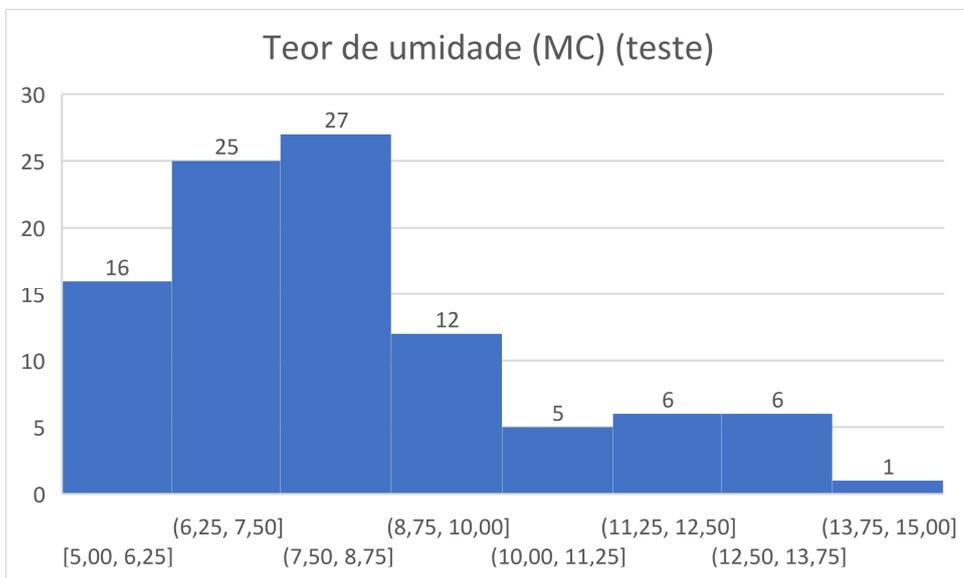
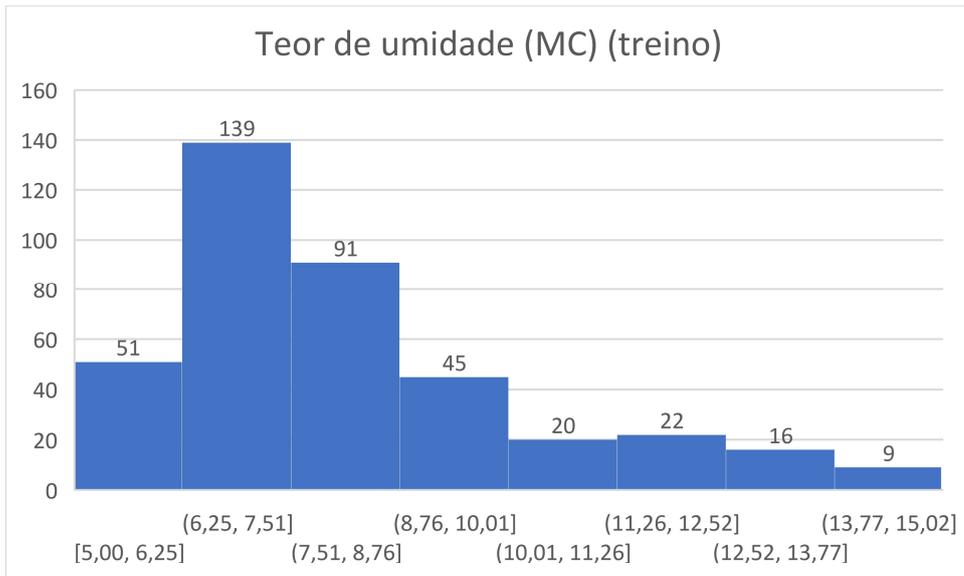


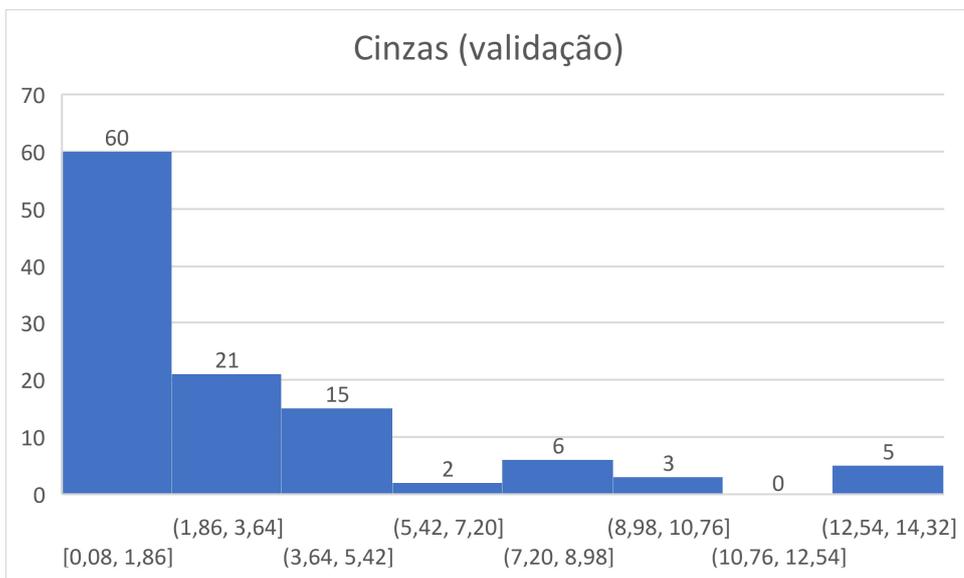
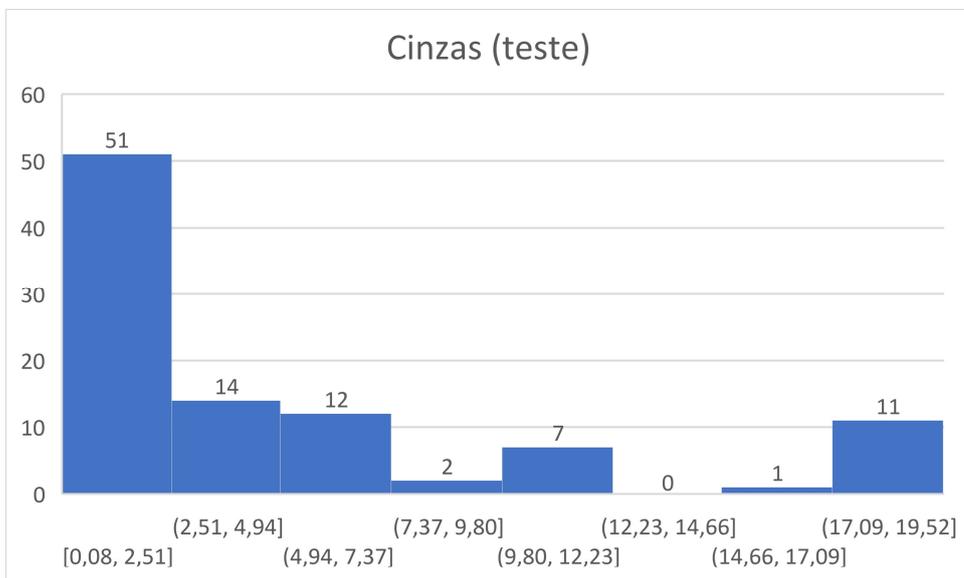
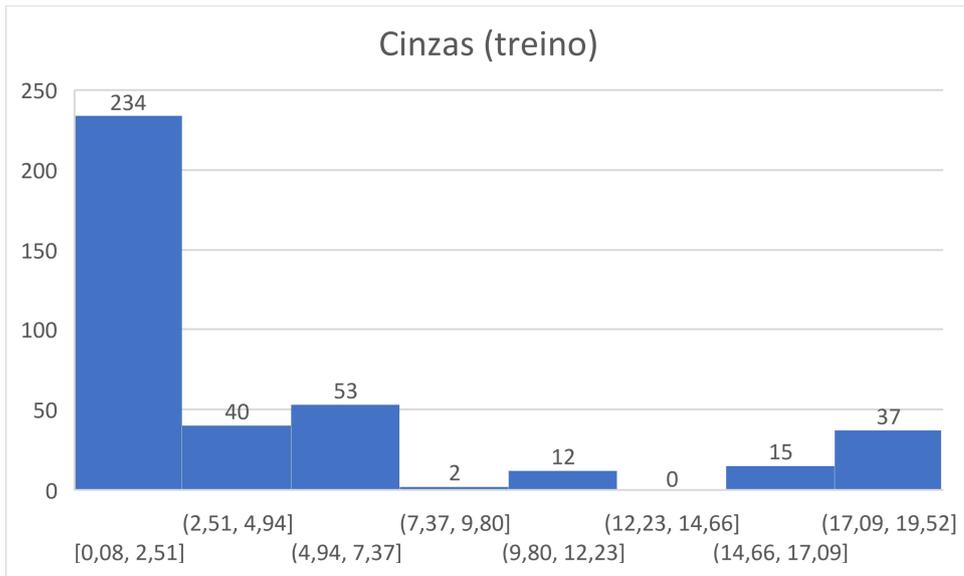


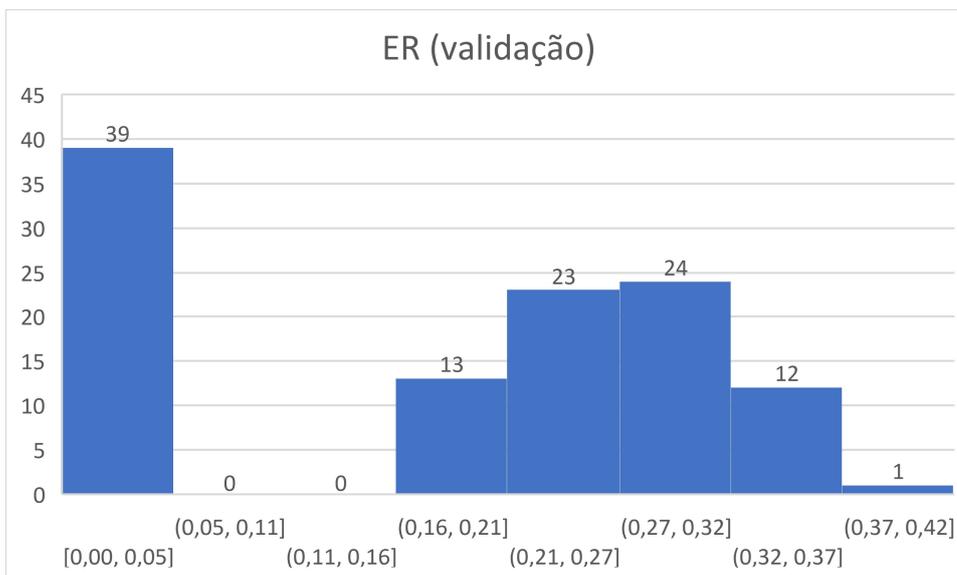
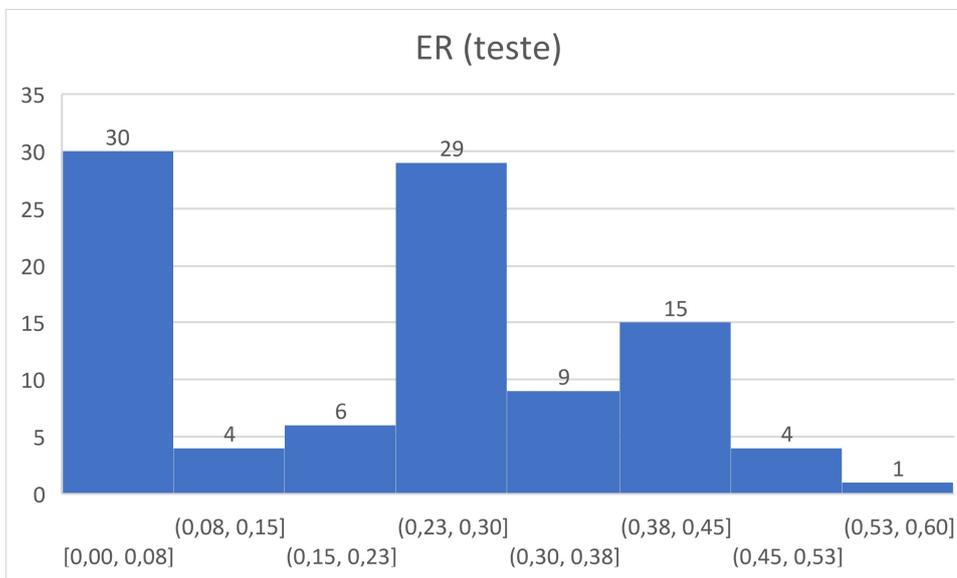
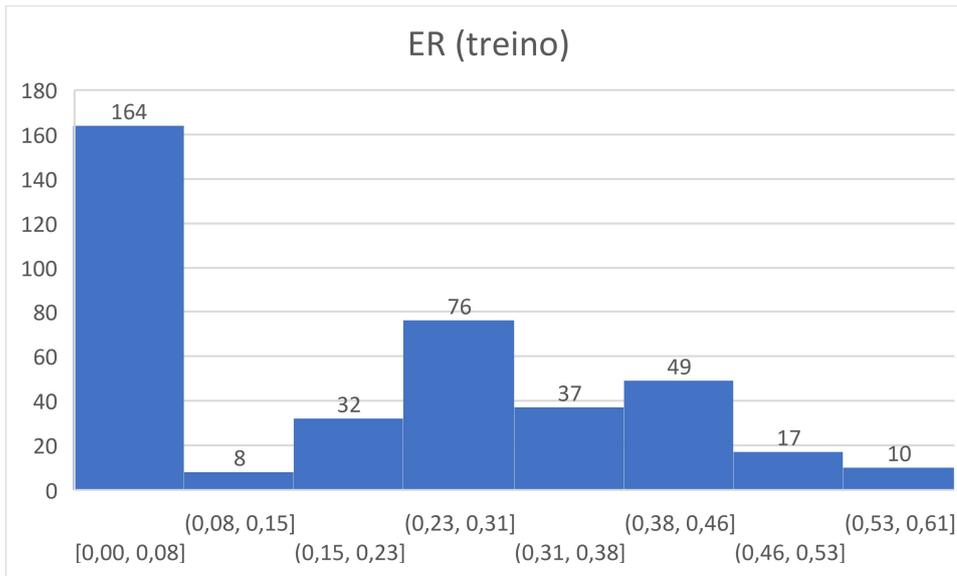


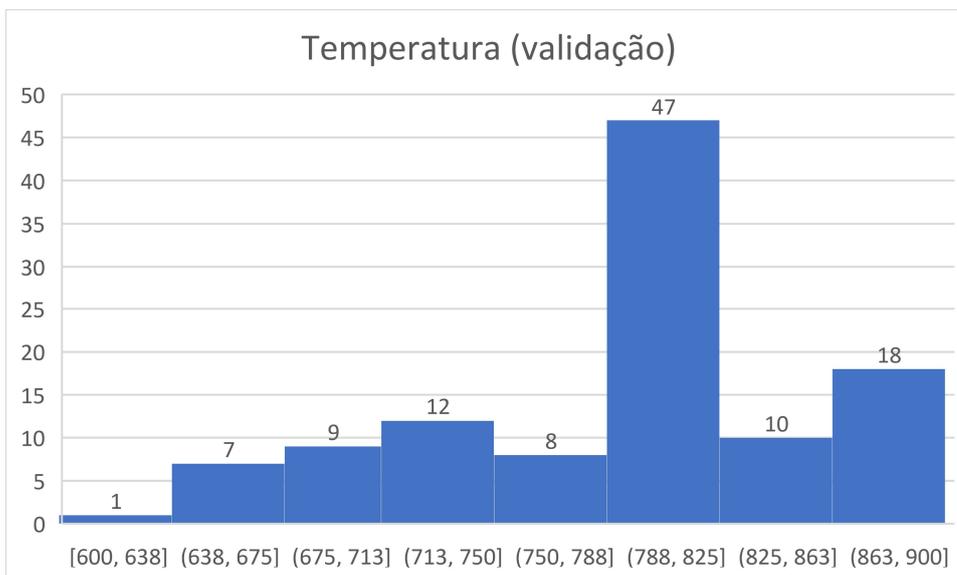
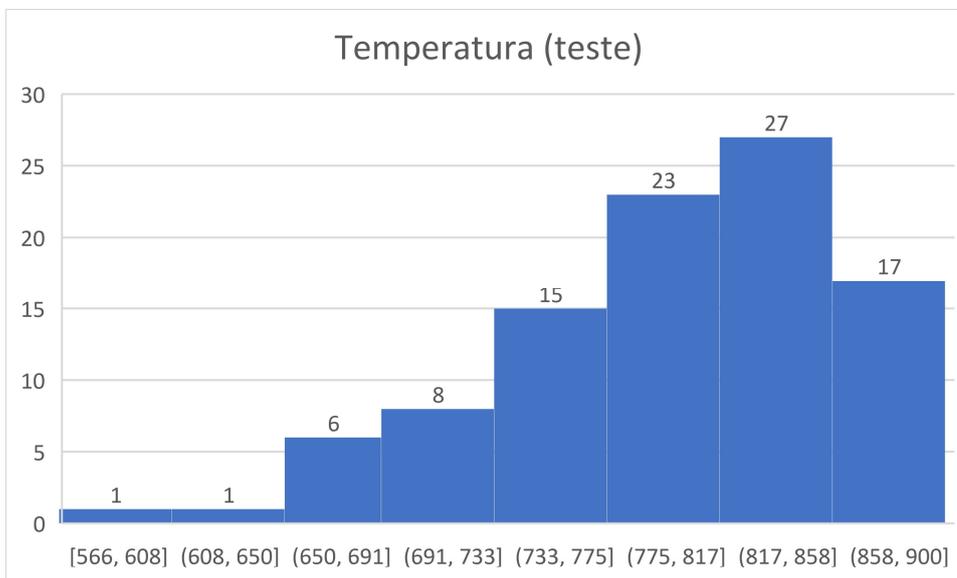
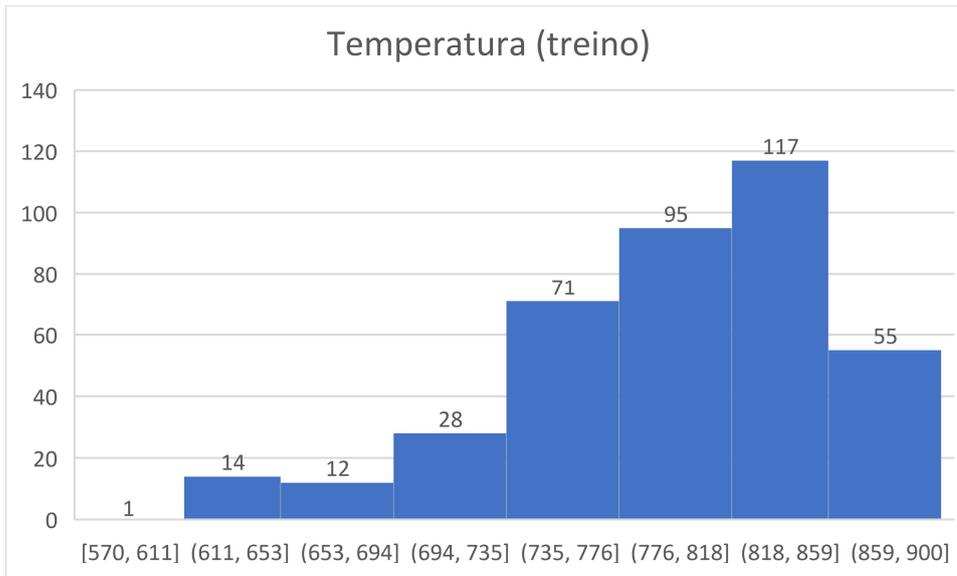


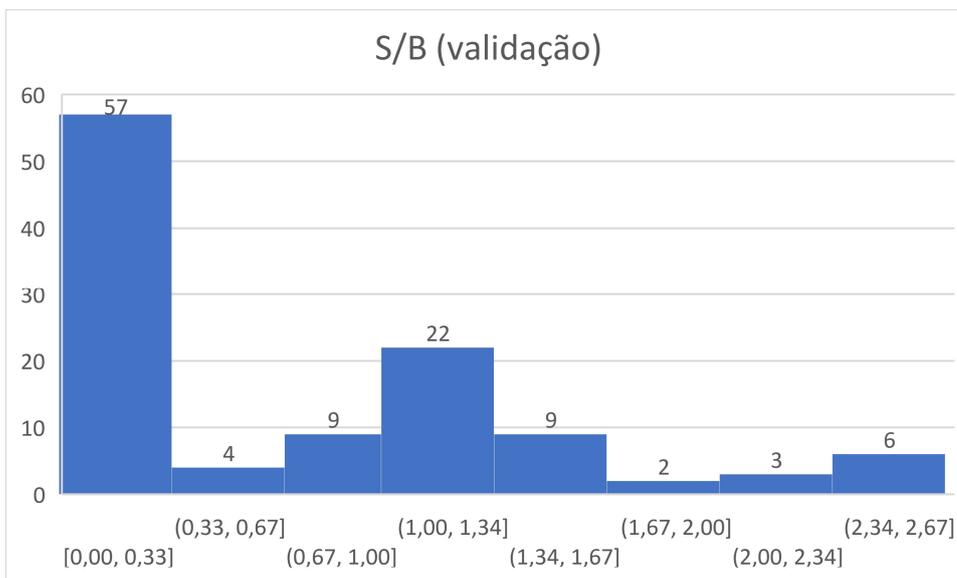
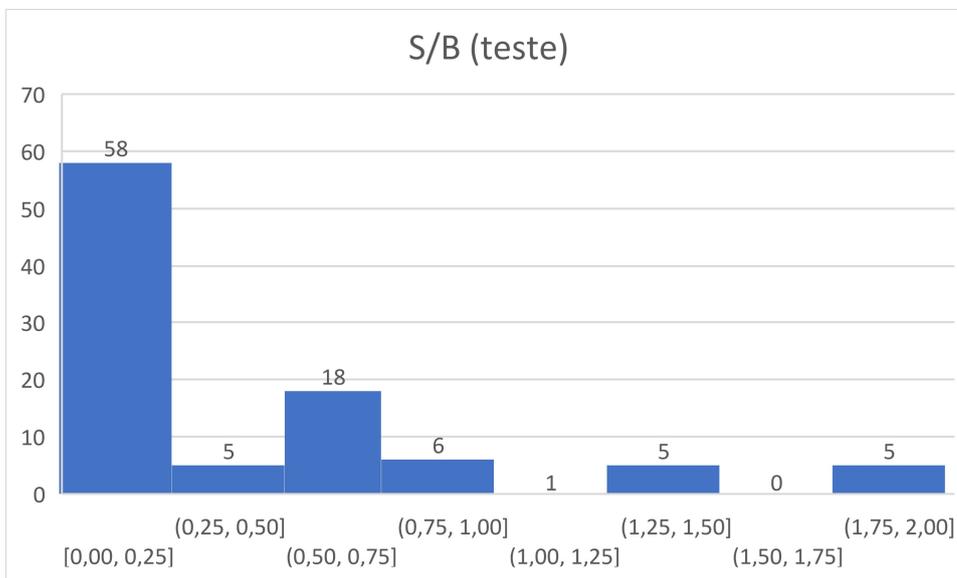
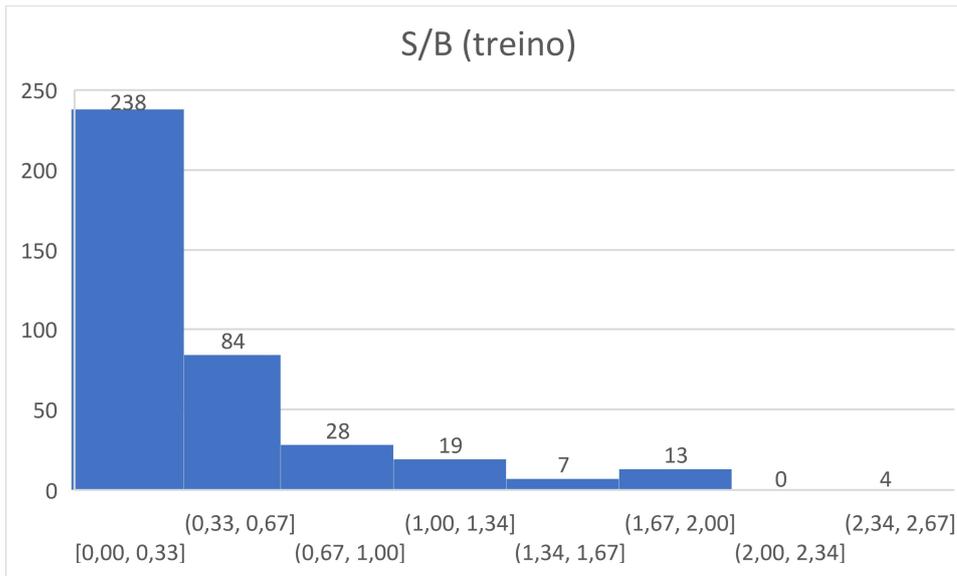












## Apêndice V – Códigos Matlab utilizados no presente trabalho

### - Primeira parte

```
clear all;
clc;

%Prepara os dados de treino para ajuste do modelo

M = load('dataset2.dat'); %carrega o arquivo de dados%
M = M'; %calcula a matriz transposta%
[div_treino,div_teste,testInd] = dividerand(M,0.8,0.2,0.0); %para divisão do banco
em treino, teste e validação
div_treino = div_treino';
div_teste = div_teste';
```

### - RNA principal

```
clear all;
clc;

%Prepara os dados de treino para ajuste do modelo

M = load('datatreino2.dat'); %carrega o arquivo de dados%
M = M'; %calcula a matriz transposta%
entrada = M(1:10,:); %define os dados de entrada da rede%
saida = M(11:14,:); %define os dados de saída da rede%
[entradan,minentrada,maxentrada,saidan,minsaida,maxsaida]=premnmx(entrada,s
aida);
%define os parâmetros máximos e mínimos das matrizes de entrada e saída e faz
normalização%

%%
%Prepara do treinamento da RNA

net = newff(minmax(entradan(:,:)),[20,20,4],{'logsig','tansig','purelin'},'trainbr');
%cria a rede, definindo número de neurônios e função de ativação das camadas
intermediária e de saída%
net.trainParam.epochs = 100; %nº de passos%
net.trainParam.goal = 1e-4; %convergência desejada%
net.initFcn = 'initlay'; %função que inicia os pesos e bias%
net.performFcn = 'sse'; %função objetivo a ser minimizada%
net.trainParam.min_grad = 1e-100; %mínimo gradiente%
net = init(net);
```

```
[net,tr] = train(net,entradan(:,:),saidan(:,:)); %realiza o treinamento da rede%
%pesos e bias da rede determinados e guardados em 'net'%
Y = sim(net,entradan(:,:)); %simula com os dados de entrada do
X = postmmx(Y,minsaida,maxsaida); %desnormaliza dados de saida%
```

```
%Mostra os pesos e bias da RNA
```

```
P1=net.IW{1}
P2=net.LW{2}
P3=net.LW{6}
```

```
B1=net.b{1}
B2=net.b{2}
B3=net.b{3}
```

```
%%
```

```
%Análise do modelo treinado com dados do treino
```

```
%Quantitativo
```

```
sse_trei1=sum(X(1,:)-saida(1,:)).^2
sse_trei2=sum(X(2,:)-saida(2,:)).^2
sse_trei3=sum(X(3,:)-saida(3,:)).^2
sse_trei4=sum(X(4,:)-saida(4,:)).^2
```

```
figure(1); %treino saida 1
```

```
[m,b,r]=postreg(X(1,:),saida(1,:)); %gráfico da saída real versus calculada%
r2_treino1=r^2; %saida 1; treino
```

```
figure(2); %treino saida 2
```

```
[m1,b1,r1]=postreg(X(2,:),saida(2,:));
r2_treino2=r1^2;
```

```
figure(3); %treino saida 3
```

```
[m2,b2,r2]=postreg(X(3,:),saida(3,:));
r2_treino3=r2^2;
```

```
figure(4); %treino saida 4
```

```
[m3,b3,r3]=postreg(X(4,:),saida(4,:));
r2_treino4=r3^2;
```

```
%Qualitativo
```

```
figure(5); %treino saida 1
```

```
plot(saida(1,:), 'or') %grafica a saida real do arquivo de treinamento%
hold on
```

```
plot(X(1,:), '-k'); %grafica a saida calculada pela rede%
```

```
xlabel('Sample','FontSize',15); %rótulo eixo x, fonte tamanho
```

```
ylabel('H2','FontSize',15); %rótulo eixo y, fonte tamanho 20%
```

```
legend('saída real','saída predita','northeast'); %legenda no canto superior direito%
```

hold off

```
figure(6); %treino saida 2
plot(saida(2,:), 'or') %grafica a saida real do arquivo de treinamento%
hold on
plot(X(2,:), '-k'); %grafica a saida calculada pela rede%
xlabel('Sample', 'FontSize', 15); %rótulo eixo x, fonte tamanho
ylabel('CO', 'FontSize', 15); %rótulo eixo y, fonte tamanho 20%
legend('saída real', 'saída predita', 'northeast'); %legenda no canto superior direito%
hold off
```

```
figure(7); %treino saida 3
plot(saida(3,:), 'or') %grafica a saida real do arquivo de treinamento%
hold on
plot(X(3,:), '-k'); %grafica a saida calculada pela rede%
xlabel('Sample', 'FontSize', 15); %rótulo eixo x, fonte tamanho
ylabel('CO2', 'FontSize', 15); %rótulo eixo y, fonte tamanho 20%
legend('saída real', 'saída predita', 'northeast'); %legenda no canto superior direito%
hold off
```

```
figure(8); %treino saida 4
plot(saida(4,:), 'or') %grafica a saida real do arquivo de treinamento%
hold on
plot(X(4,:), '-k'); %grafica a saida calculada pela rede%
xlabel('Sample', 'FontSize', 15); %rótulo eixo x, fonte tamanho
ylabel('CH4', 'FontSize', 15); %rótulo eixo y, fonte tamanho 20%
legend('saída real', 'saída predita', 'northeast'); %legenda no canto superior direito%
hold off
```

```
save '20_20_logsig_tansig_purelin_trainbr_3' net; %salva a rede criada%
```

```
%%
```

```
%Testa o modelo de RNA treinado
```

```
%Prepara os dados de teste para simulação
```

```
N = load('datatest2.dat'); %carrega arquivo de dados para teste%
```

```
N = N';
```

```
in = N(1:10,:); %dados de entrada da rede%
```

```
out = N(11:14,:); %dados de saida da rede%
```

```
[inn] = tramnmx(in, minentrada, maxentrada); %normaliza dados%
```

```
Yt = sim(net, inn(:, :)); %simula com os dados de entrada do arquivo de teste%
```

```
Xt = postmnmx(Yt, minsaida, maxsaida); %desnormaliza dados de saida%
```

```
%Análise do modelo treinado com dados do teste (validação com dados diferentes
```

```
%Quantitativo
```

```
sse_test1 = sum(Xt(1,:) - out(1,:)).^2
```

```
sse_test2 = sum(Xt(2,:) - out(2,:)).^2
```

```
sse_test3=sum(Xt(3,:)-out(3,:)).^2
sse_test4=sum(Xt(4,:)-out(4,:)).^2
```

```
figure(9); %teste saida 1
[m4,b4,r4]=postreg(Xt(1,:),out(1,:)); %grafico da saida real versus calculada%
r2_teste1=r4^2;
```

```
figure(10); %teste saida 2
[m5,b5,r5]=postreg(Xt(2,:),out(2,:));
r2_teste2=r5^2;
```

```
figure(11); %teste saida 3
[m6,b6,r6]=postreg(Xt(3,:),out(3,:));
r2_teste3=r6^2;
```

```
figure(12); %teste saida 4
[m7,b7,r7]=postreg(Xt(4,:),out(4,:));
r2_teste4=r7^2;
```

```
%Qualitativo
```

```
figure(13); %teste saida 1
plot(out(1,:), 'or') %grafica a saida real do arquivo de teste%
hold on
plot(Xt(1,:), '-k'); %grafica a saida calculada pela rede%
xlabel('Sample','FontSize',15); %rótulo eixo x, fonte tamanho 20%
ylabel('H2','FontSize',15); %rótulo eixo y, fonte tamanho 20%
legend('saída real','saída predita','northeast'); %legenda no canto superior direito%
colordef white;
hold off
```

```
figure(14); %teste saida 2
plot(out(2,:), 'or') %grafica a saida real do arquivo de teste%
hold on
plot(Xt(2,:), '-k'); %grafica a saida calculada pela rede%
xlabel('Sample','FontSize',15); %rótulo eixo x, fonte tamanho 20%
ylabel('CO','FontSize',15); %rótulo eixo y, fonte tamanho 20%
legend('saída real','saída predita','northeast'); %legenda no canto superior direito%
colordef white;
hold off
```

```
figure(15); %teste saida 3
plot(out(3,:), 'or') %grafica a saida real do arquivo de teste%
hold on
plot(Xt(3,:), '-k'); %grafica a saida calculada pela rede%
xlabel('Sample','FontSize',15); %rótulo eixo x, fonte tamanho 20%
ylabel('CO2','FontSize',15); %rótulo eixo y, fonte tamanho 20%
legend('saída real','saída predita','northeast'); %legenda no canto superior direito%
colordef white;
hold off
```

```

figure(16); %teste saida 4
plot(out(4,:), 'or') %grafica a saida real do arquivo de teste%
hold on
plot(Xt(4,:), '-k'); %grafica a saida calculada pela rede%
xlabel('Sample','FontSize',15); %rótulo eixo x, fonte tamanho 20%
ylabel('CH4','FontSize',15); %rótulo eixo y, fonte tamanho 20%
legend('saída real','saída predita','northeast'); %legenda no canto superior direito%
colordef white;
hold off

```

### - Superfícies de resposta

```
% SCRIPT GOAL: to analyze the response surfaces for best ANN
```

```

close all;
clear all;
clc;

```

```
% Load ANN
```

```

ANN = load('10_10_logsig_logsig_purelin_trainbr_rede6.mat');
ANN = ANN.net;

```

```
% Load training database for normalization
```

```

M = load('dadostreino.dat'); %carrega o arquivo de dados%
M = M'; %calcula a matriz transposta%
entrada = M(1:10,:); %define os dados de entrada da rede%
saida = M(11:14,:); %define os dados de saída da rede%
[entradan,minentrada,maxentrada,saidan,minsaida,maxsaida]=premnmx(entrada,s
aida);
%define os parâmetros máximos e mínimos das matrizes de entrada e
%saída e faz normalização%

```

```
% Average values for fixed variables
```

```

average_values = zeros(10,1);
average_values(1) = 48.99; % Input 1
average_values(2) = 5.95; % Input 2
average_values(3) = 44.39; % Input 3
average_values(4) = 0.53; % Input 4
average_values(5) = 75.51; % Input 5
average_values(6) = 8.29; % Input 6
average_values(7) = 4.35; % Input 7
average_values(8) = 0.19; % Input 8
average_values(9) = 799; % Input 9
average_values(10) = 0.39; % Input 10

```

```
% Desired plots
```

```

desired_plots = {[6 9]};

% Variables limits
limits_vec = zeros(10,2);
limits_vec(1,:) = [38.92 56.58];%C
limits_vec(2,:) = [3.6 7.82]; %H
limits_vec(3,:) = [37.04 53.69]; %O
limits_vec(4,:) = [0 3.1]; %N
limits_vec(5,:) = [55.54 94.13]; %VM
limits_vec(6,:) = [5 15.02]; %MC
limits_vec(7,:) = [0.08 19.52]; %ash
limits_vec(8,:) = [0 0.61]; %ER
limits_vec(9,:) = [570 900]; %T
limits_vec(10,:) = [0 2.67]; %S/B

% Input vector
inputs = trammx(average_values,minentrada,maxentrada);

% Plot intervals
intervals = {0,0,0,0,0,0,0,0,0,0}; % será preenchido abaixo para variáveis
contínuas

num_pts = 50; % total de pontos para variáveis contínuas

for i=1:10
    delta = (limits_vec(i,2)-limits_vec(i,1))/(num_pts-1);
    intervals{i} = (limits_vec(i,1):delta:limits_vec(i,2)).';
end

% intervals normalization
intervals_norm = {0,0,0,0,0,0,0,0,0,0};

for i=1:10
    intervals_norm{i} = trammx(intervals{i},minentrada(i),maxentrada(i));
end

% Variables txt
axis_txt = {'Carbono', 'Hidrogênio', 'Oxigênio', 'Nitrogênio', 'Material volátil',
'Teor de umidade (%)', 'Cinzas', 'ER', 'Temperatura (°C)', 'S/B'};
saida_txt = {'H2', 'CO', 'CO2', 'CH4'};
titulo_txt = {'Superfície de resposta para saída 1', 'Superfície de resposta para saída
2', 'Superfície de resposta para saída 3', 'Superfície de resposta para saída 4'};

% Plots
for i = 1:size(desired_plots,2) %PARA CADA DUPLA NA LISTA desired_plots

    variable_1 = desired_plots{i}(1);
    variable_2 = desired_plots{i}(2);
    [X,Y] = meshgrid(intervals_norm{variable_1},intervals_norm{variable_2});
    size_var_1 = size(intervals_norm{variable_1},1);

```

```

size_var_2 = size(intervals_norm{variable_2},1);
Z = {};
Z_desnorm = {};

for t = 1:size_var_1
    for u = 1:size_var_2
        inputs(variable_1) = X(u,t);
        inputs(variable_2) = Y(u,t);
        Z{u,t} = sim(ANN,inputs);
        Z_desnorm{u,t} = postmmx(Z{u,t},minsaida,maxsaida);
    end
end

```

```

variable_1_desnorm =
postmmx(X,minentrada(variable_1),maxentrada(variable_1));
variable_2_desnorm =
postmmx(Y,minentrada(variable_2),maxentrada(variable_2));

```

```

fig = figure(1);
Z_desnormj = zeros();

```

```

for t = 1:size_var_1
    for u = 1:size_var_2
        Z_desnormj(u,t) = Z_desnorm{u,t}(1);
        if(Z_desnormj(u,t) < 0)
            Z_desnormj(u,t) = 0;
        end
    end
end
end

```

```

surf(variable_1_desnorm,variable_2_desnorm,Z_desnormj)
xlabel(axis_txt{variable_1});
ylabel(axis_txt{variable_2});
zlabel(saida_txt{1});

```

```

title(titulo_txt{1});

```

```

saveas(fig, strcat('Superficie_de_resposta_',saida_txt{1},'_var',string(variable_1),string(variable_2)), 'fig');

```

```

saveas(fig, strcat('Superficie_de_resposta_',saida_txt{1},'_var',string(variable_1),string(variable_2)), 'png');
close(fig);

```

```

fig = figure(2);
Z_desnormj = zeros();

```

```

for t = 1:size_var_1

```

```

for u = 1:size_var_2
    Z_desnormj(u,t) = Z_desnorm{u,t}(2);
    if(Z_desnormj(u,t) < 0)
        Z_desnormj(u,t) = 0;
    end
end
end

surf(variable_1_desnorm,variable_2_desnorm,Z_desnormj)
xlabel(axis_txt{variable_1});
ylabel(axis_txt{variable_2});
zlabel(saida_txt{2});

title(titulo_txt{2});

saveas(fig,strcat('Superficie_de_resposta_',saida_txt{2},'var',string(variable_1),string(variable_2)),'fig');

saveas(fig,strcat('Superficie_de_resposta_',saida_txt{2},'var',string(variable_1),string(variable_2)),'png');
close(fig);

fig = figure(3);
Z_desnormj = zeros();

for t = 1:size_var_1
    for u = 1:size_var_2
        Z_desnormj(u,t) = Z_desnorm{u,t}(3);
        if(Z_desnormj(u,t) < 0)
            Z_desnormj(u,t) = 0;
        end
    end
end

surf(variable_1_desnorm,variable_2_desnorm,Z_desnormj)
xlabel(axis_txt{variable_1});
ylabel(axis_txt{variable_2});
zlabel(saida_txt{3});

title(titulo_txt{3});

saveas(fig,strcat('Superficie_de_resposta_',saida_txt{3},'var',string(variable_1),string(variable_2)),'fig');

saveas(fig,strcat('Superficie_de_resposta_',saida_txt{3},'var',string(variable_1),string(variable_2)),'png');
close(fig);

```

```

fig = figure(4);
Z_desnormj = zeros();

for t = 1:size_var_1
    for u = 1:size_var_2
        Z_desnormj(u,t) = Z_desnorm{u,t}(4);
        if(Z_desnormj(u,t) < 0)
            Z_desnormj(u,t) = 0;
        end
    end
end

surf(variable_1_desnorm,variable_2_desnorm,Z_desnormj)
xlabel(axis_txt{variable_1});
ylabel(axis_txt{variable_2});
zlabel(saida_txt{4});

title(titulo_txt{4});

saveas(fig, strcat('Superficie_de_resposta_',saida_txt{4}, 'var', string(variable_1), string(variable_2)), 'fig');

saveas(fig, strcat('Superficie_de_resposta_',saida_txt{4}, 'var', string(variable_1), string(variable_2)), 'png');
close(fig);

inputs = trammx(average_values,minentrada,maxentrada); % resets inputs for future iterations

end

```

### - Validação RNA

```

clear all;
clc;

%%
%Validação do modelo de RNA – rede 6

%Carregar a rede pré-definida
load('10_10_logsig_logsig_purelin_trainbr_rede6.mat','net');

%Prepara os dados de validação para simulação

M = load('dadostreino.dat'); %carrega o arquivo de dados%
M = M'; %calcula a matriz transposta%
entrada = M(1:10,:); %define os dados de entrada da rede%

```

```

saida = M(11:14,:); %define os dados de saída da rede%
[entradan,minentrada,maxentrada,saidan,minsaida,maxsaida]=premmx(entrada,s
aida);
%define os parâmetros máximos e mínimos das matrizes de entrada e
%saída e faz normalização%

N = load('dadosval.dat'); %carrega arquivo de dados de validação%
N = N';
in = N(1:10,:); %dados de entrada da rede%
out = N(11:14,:); %dados de saída da rede%
[inn] = tramnmx(in,minentrada,maxentrada); %normaliza dados%
Yt = sim(net,inn(:,:)); %simula com os dados de entrada do arquivo
%de teste% saída predita normalizada
Ydes = postmmx(Yt,minsaida,maxsaida); %desnormaliza dados de saída para
gerar gráfico%

```

```
%Análise do modelo validado
```

```
%Quantitativo
```

```

sse_val1=sum(Ydes(1,:)-out(1,:)).^2;
sse_val2=sum(Ydes(2,:)-out(2,:)).^2;
sse_val3=sum(Ydes(3,:)-out(3,:)).^2;
sse_val4=sum(Ydes(4,:)-out(4,:)).^2;

```

```
figure(17); %validação saída 1
```

```

[m1,b1,r1]=postreg(Ydes(1,:),out(1,:)); %grafico da saída real versus calculada%
r2_val1=r1^2;

```

```
figure(18); % validação saída 2
```

```

[m2,b2,r2]=postreg(Ydes(2,:),out(2,:));
r2_val2=r2^2;

```

```
figure(19); % validação saída 3
```

```

[m3,b3,r3]=postreg(Ydes(3,:),out(3,:));
r2_val3=r3^2;

```

```
figure(20); % validação saída 4
```

```

[m4,b4,r4]=postreg(Ydes(4,:),out(4,:));
r2_val4=r4^2;

```

```
%Qualitativo
```

```
figure(21); % validação saída 1
```

```

plot(out(1,:),'or') %grafica a saída real do arquivo de validação%
hold on

```

```
plot(Ydes(1,:),'-k'); %grafica a saída calculada pela rede%
```

```
xlabel('Sample','FontSize',15); %rótulo eixo x, fonte tamanho 20%
```

```
ylabel('H2','FontSize',15); %rótulo eixo y, fonte tamanho 20%
```

```
legend('saída real','saída predita','northeast'); %legenda no canto
```

```

%superior direito%
colordef white;
hold off

figure(22); % validação saída 2
plot(out(2,:), 'or') %grafica a saída real do arquivo de validação%
hold on
plot(Ydes(2,:), '-k'); %grafica a saída calculada pela rede%
xlabel('Sample', 'FontSize', 15); %rótulo eixo x, fonte tamanho
%20%
ylabel('CO', 'FontSize', 15); %rótulo eixo y, fonte tamanho 20%
legend('saída real', 'saída predita', 'northeast'); %legenda no canto
%superior direito%
colordef white;
hold off

figure(23); % validação saída 3
plot(out(3,:), 'or') %grafica a saída real do arquivo de validação%
hold on
plot(Ydes(3,:), '-k'); %grafica a saída calculada pela rede%
xlabel('Sample', 'FontSize', 15); %rótulo eixo x, fonte tamanho
%20%
ylabel('CO2', 'FontSize', 15); %rótulo eixo y, fonte tamanho 20%
legend('saída real', 'saída predita', 'northeast'); %legenda no canto
%superior direito%
colordef white;
hold off

figure(24); % validação saída 4
plot(out(4,:), 'or') %grafica a saída real do arquivo de validação%
hold on
plot(Ydes(4,:), '-k'); %grafica a saída calculada pela rede%
xlabel('Sample', 'FontSize', 15); %rótulo eixo x, fonte tamanho
%20%
ylabel('CH4', 'FontSize', 15); %rótulo eixo y, fonte tamanho 20%
legend('saída real', 'saída predita', 'northeast'); %legenda no canto
%superior direito%
colordef white;
hold off

```

### - Matriz para análise quantitativa

```

close all;
clear all;
clc;

%%Carregar BD do treino e normalizar
M = load('dadostreino.dat'); %carrega o arquivo de dados%

```

```

M = M'; %calcula a matriz transposta%
entrada = M(1:10,:); %define os dados de entrada da rede%
saida = M(11:14,:); %define os dados de saída da rede%
[entradan,minentrada,maxentrada,saidan,minsaida,maxsaida]=premnmx(entrada,s
aida);
%define os parâmetros máximos e mínimos das matrizes de entrada e
%saída e faz normalização%

%%Carregar BD da validação e normalizar com mesmo máx e min do treino
N = load('dados2val.dat'); %carrega arquivo de dados de validação%
N = N';
in = N(1:10,:); %dados de entrada da rede%
out = N(11:14,:); %dados de saída da rede%
[inn] = tramnmx(in,minentrada,maxentrada); %normaliza dados%

%Distância de treinamento e validação do ponto 1 até ponto 1
%disp(sqrt(sum((entradan(:,1) - inn(:,1)).^2)))

%%Matriz com somas das distâncias ponto a ponto

%Matriz treino (linha) com validação (coluna)
matriz_val = zeros(393,112);

for i = 1:393 %treino
    for j = 1:112 %val

        matriz_val(i,j) = sqrt(sum((entradan(:,i) - inn(:,j)).^2));

    end
end

%%Ordenar cada coluna da matriz_val em ordem crescente
B = sort(matriz_val);

```

## Apêndice VI – Topologias das 33 redes neurais artificiais testadas

Teste	Topologia da rede									
	Número de neurônios				Função de ativação					Algoritmo de treinamento
	1ª CI	2ª CI	3ª CI	4ª CI	1ª CI	2ª CI	3ª CI	4ª CI	CF	
1	10	N/A	N/A	N/A	logsig	N/A	N/A	N/A	purelin	trainbr
2	10	N/A	N/A	N/A	tansig	N/A	N/A	N/A	purelin	trainbr
3	10	N/A	N/A	N/A	tansig	N/A	N/A	N/A	purelin	trainlm
4	20	N/A	N/A	N/A	tansig	N/A	N/A	N/A	purelin	trainbr
5	10	10	N/A	N/A	logsig	tansig	N/A	N/A	purelin	trainbr
6	10	10	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr
7	10	10	N/A	N/A	tansig	tansig	N/A	N/A	purelin	trainbr
8	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr
9	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	traingdx
10	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	traingd
11	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	traingda
12	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	traingdm
13	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbfg
14	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	traincgb
15	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	traincgf
16	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	traincgp
17	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainrp
18	20	20	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainlm
19	40	40	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr
20	60	60	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr
21	40	40	40	N/A	logsig	logsig	tansig	N/A	purelin	trainbr
22	30	30	30	30	logsig	logsig	tansig	tansig	purelin	trainbr
23	20	60	N/A	N/A	tansig	logsig	N/A	N/A	purelin	trainbr
24	60	80	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr
25	60	80	N/A	N/A	logsig	logsig	N/A	N/A	tansig	trainbr
26	60	80	N/A	N/A	logsig	logsig	N/A	N/A	satlin	trainbr
27	10	N/A	N/A	N/A	tansig	N/A	N/A	N/A	satlin	trainbr
28	100	N/A	N/A	N/A	tansig	N/A	N/A	N/A	purelin	trainbr
29	100	N/A	N/A	N/A	logsig	N/A	N/A	N/A	purelin	trainbr
30	80	80	N/A	N/A	logsig	logsig	N/A	N/A	purelin	trainbr
31	200	N/A	N/A	N/A	tansig	N/A	N/A	N/A	purelin	trainbr
32	400	N/A	N/A	N/A	tansig	N/A	N/A	N/A	purelin	trainbr
33	400	N/A	N/A	N/A	tansig	N/A	N/A	N/A	purelin	trainlm

CI: camada intermediária, CF: camada final, N/A: não aplicável

## Apêndice VII – Análise quantitativa das 33 redes neurais artificiais testadas

Teste	Saída 1 (H2)						Saída 2 (CO)					
	R <sup>2</sup> treino1	R <sup>2</sup> teste1	Média ponderada R <sup>2</sup>	SSE treino1	SSE teste1	Média ponderada SSE	R <sup>2</sup> treino2	R <sup>2</sup> teste2	Média ponderada R <sup>2</sup>	SSE treino2	SSE teste2	Média ponderada SSE
1	0,9486	0,8929	0,9375	130,7826	2,29E+03	561,9262	0,7871	0,5456	0,7389	5,8788	5,01E+03	1004,6447
2	0,9316	0,8887	0,9230	0,2663	113,9837	22,9635	0,7187	0,6716	0,7093	0,0683	4,63E+03	924,9671
3	0,8993	0,8359	0,8866	0,0404	2,10E+04	4195,6698	0,6751	0,6925	0,6786	0,0168	6,43E+03	1283,1947
4	0,9794	0,8857	0,9607	6,5926	1,78E+03	360,3322	0,8939	0,5246	0,8202	3,9158	1,08E+04	2164,7228
5	0,9747	0,8895	0,9577	2,86E+03	1,87E+03	2660,4802	0,8890	0,6125	0,8338	356,6102	1,85E+04	3980,8917
6	0,9755	0,9119	0,9628	0,2880	1,11E+03	221,3792	0,8857	0,7003	0,8487	0,6671	8,37E+03	1671,0847
7	0,9760	0,8563	0,9521	0,0163	5,18E+03	1032,9051	0,8799	0,5791	0,8199	0,5118	1,14E+04	2272,3730
8	0,9838	0,9491	0,9769	0,2426	3,97E+03	792,9563	0,9033	0,7564	0,8740	0,6819	4,54E+03	906,5368
9	0,8180	0,8121	0,8168				0,4635	0,4387	0,4586			
10	0,0527	0,1273	0,0676				0,0340	0,0557	0,0383			
11	0,6298	0,5692	0,6177				0,2642	0,1497	0,2413			
12	5,81E-57	5,01E-57	0,0000				3,15E-55	5,24E-56	0,0000			
13	0,8167	0,7553	0,8044				0,3044	0,2093	0,2854			
14	0,9151	0,9056	0,9132				0,6330	0,4889	0,6042			
15	0,8672	0,8326	0,8603				0,4964	0,2795	0,4531			
16	0,8910	0,8581	0,8844				0,6289	0,5839	0,6199			
17	0,9394	0,8894	0,9294	461,0729	2,19E+03	805,9545	0,7113	0,6169	0,6925	8,9836	1,40E+04	2794,9013
18	0,9866	0,3855	0,8666	0,1287	5,97E+04	1,19E+04	0,9307	0,2563	0,7961	0,3007	3,22E+04	6418,5421
19	0,9846	0,9343	0,9746	1,4509	4,23E+03	846,2766	0,9086	0,6893	0,8648	3,4182	4,56E+03	912,7188
20	0,9835	0,9286	0,9725	50,4801	43,3851	49,0640	0,9003	0,7168	0,8637	51,2752	5,38E+03	1115,2687
21	0,9847	0,9508	0,9779	4,6858	49,0948	13,5495	0,9058	0,7244	0,8696	1,9109	7,28E+03	1454,6639
22	0,9805	0,9168	0,9678	5,8950	4,73E+03	947,9534	0,8871	0,7150	0,8528	0,0523	2,14E+03	426,5514
23	0,9839	0,8765	0,9625	1,5225	2,53E+03	507,0463	0,9078	0,6375	0,8539	0,4969	2,91E+03	581,2323

Teste	Saída 1 (H2)						Saída 2 (CO)					
	R <sup>2</sup> treino1	R <sup>2</sup> teste1	Média ponderada R <sup>2</sup>	SSE treino1	SSE teste1	Média ponderada SSE	R <sup>2</sup> treino2	R <sup>2</sup> teste2	Média ponderada R <sup>2</sup>	SSE treino2	SSE teste2	Média ponderada SSE
24	0,9847	0,9548	<b>0,9787</b>	10,1861	494,4861	<b>106,8488</b>	0,9080	0,7640	<b>0,8793</b>	193,7042	4,95E+03	<b>1143,1058</b>
25	0,9821	0,9200	<b>0,9697</b>	9,94E+03	835,5214	<b>8122,8128</b>	0,8937	0,7431	<b>0,8636</b>	1,55E+03	7,33E+03	<b>2703,3652</b>
26	0,0000	8,02E-56	<b>0,0000</b>			<b>0,0000</b>	3,18E-54	2,29E-56	<b>0,0000</b>			<b>0,0000</b>
27	0,4346	0,4575	<b>0,4392</b>			<b>0,0000</b>	3,18E-54	2,29E-56	<b>0,0000</b>			<b>0,0000</b>
28	0,9835	0,9264	<b>0,9721</b>	0,0280	745,4016	<b>148,7991</b>	0,9009	0,6162	<b>0,8441</b>	4,4157	6,31E+03	<b>1262,9841</b>
29	0,9822	0,9046	<b>0,9667</b>	7,1704	31,3178	<b>11,9900</b>	0,9000	0,5815	<b>0,8364</b>	0,1488	7,98E+03	<b>1593,6270</b>
30	0,9841	0,9330	<b>0,9739</b>	1,4810	1,33E+03	<b>266,9630</b>	0,9040	0,6364	<b>0,8506</b>	0,4590	7,29E+03	<b>1456,3161</b>
31	0,9826	0,9201	<b>0,9701</b>	345,1554	572,2621	<b>390,4842</b>	0,8932	0,6565	<b>0,8460</b>	14,3255	5,70E+03	<b>1149,8430</b>
32	0,9828	0,9087	<b>0,9680</b>	8,7471	3,07E+03	<b>619,0522</b>	0,8993	0,6257	<b>0,8447</b>	0,5306	5,20E+03	<b>1038,5062</b>
33	0,9876	0,2155	<b>0,8335</b>	0,1759	2,00E+05	<b>3,99E+04</b>	0,9362	0,0327	<b>0,7559</b>	0,6417	1,65E+04	<b>3298,3833</b>

Teste	Saída 3 (CO2)						Saída 4 (CH4)					
	R <sup>2</sup> treino3	R <sup>2</sup> teste3	Média ponderada R <sup>2</sup>	SSE treino3	SSE teste3	Média ponderada SSE	R <sup>2</sup> treino4	R <sup>2</sup> teste4	Média ponderada R <sup>2</sup>	SSE treino4	SSE teste4	Média ponderada SSE
1	0,9037	0,7583	<b>0,8747</b>	48,9082	32,7026	<b>45,6737</b>	0,8675	0,8686	<b>0,8677</b>	12,1094	140,7543	<b>37,7860</b>
2	0,8982	0,7839	<b>0,8754</b>	0,3303	1,12E+03	<b>223,9678</b>	0,8607	0,8640	<b>0,8614</b>	0,0018	40,5855	<b>8,1020</b>
3	0,9046	0,7001	<b>0,8638</b>	0,0194	1,08E+03	<b>215,6554</b>	0,8303	0,8223	<b>0,8287</b>	3,37E-05	79,2608	<b>15,8199</b>
4	0,9259	0,8012	<b>0,9010</b>	0,3263	1,78E+03	<b>356,4143</b>	0,9035	0,8856	<b>0,8999</b>	1,0889	307,6548	<b>62,2772</b>
5	0,9243	0,7722	<b>0,8939</b>	7,7163	372,6630	<b>80,5570</b>	0,9003	0,8520	<b>0,8907</b>	10,2857	87,0706	<b>25,6114</b>
6	0,9274	0,8135	<b>0,9047</b>	2,6530	252,6077	<b>52,5421</b>	0,9063	0,8581	<b>0,8967</b>	7,6254	371,6670	<b>80,2854</b>
7	0,9250	0,8301	<b>0,9061</b>	0,0801	62,7691	<b>12,5924</b>	0,9063	0,8861	<b>0,9023</b>	0,2388	639,8115	<b>127,8928</b>
8	0,9307	0,8159	<b>0,9078</b>	0,2180	374,5634	<b>74,9346</b>	0,9169	0,9248	<b>0,9185</b>	0,0089	18,3337	<b>3,6664</b>
9	0,6693	0,5961	<b>0,6547</b>				0,6444	0,5874	<b>0,6330</b>			
10	0,1285	0,0648	<b>0,1158</b>				0,0159	0,0391	<b>0,0205</b>			
11	0,4145	0,4091	<b>0,4134</b>				0,4403	0,3435	<b>0,4210</b>			
12	5,88E-54	6,61E-56	<b>0,0000</b>				9,58E-59	2,26E-58	<b>0,0000</b>			
13	0,5240	0,5366	<b>0,5265</b>				0,4367	0,4325	<b>0,4359</b>			
14	0,8439	0,7093	<b>0,8170</b>				0,8191	0,8294	<b>0,8212</b>			
15	0,7743	0,6600	<b>0,7515</b>				0,7859	0,7488	<b>0,7785</b>			
16	0,8142	0,7125	<b>0,7939</b>				0,8233	0,7862	<b>0,8159</b>			
17	0,8752	0,7000	<b>0,8402</b>	143,9546	942,0302	<b>303,2446</b>	0,8513	0,7795	<b>0,8370</b>	109,1319	15,1168	<b>90,3672</b>
18	0,9671	0,0866	<b>0,7914</b>	1,0488	2,79E+05	<b>5,57E+04</b>	0,9372	0,2657	<b>0,8032</b>	0,5344	7,12E+03	<b>1,42E+03</b>
19	0,9348	0,7904	<b>0,9060</b>	0,1074	501,8162	<b>100,2448</b>	0,9203	0,9060	<b>0,9174</b>	0,2535	16,0874	<b>3,4138</b>
20	0,9290	0,8225	<b>0,9077</b>	4,6082	382,1292	<b>79,9586</b>	0,9168	0,8808	<b>0,9096</b>	0,4463	0,6926	<b>0,4955</b>
21	0,9328	0,7980	<b>0,9059</b>	23,2530	88,8060	<b>36,3369</b>	0,9184	0,9242	<b>0,9196</b>	0,8451	204,8855	<b>41,5701</b>
22	0,9269	0,7846	<b>0,8985</b>	16,9382	194,0372	<b>52,2859</b>	0,9042	0,8770	<b>0,8988</b>	0,1335	49,2148	<b>9,9298</b>
23	0,9347	0,7969	<b>0,9072</b>	0,0220	81,3173	<b>16,2479</b>	0,9196	0,9016	<b>0,9160</b>	0,6717	163,9040	<b>33,2517</b>

Teste	Saída 3 (CO2)						Saída 4 (CH4)					
	R <sup>2</sup> treino3	R <sup>2</sup> teste3	Média ponderada R <sup>2</sup>	SSE treino3	SSE teste3	Média ponderada SSE	R <sup>2</sup> treino4	R <sup>2</sup> teste4	Média ponderada R <sup>2</sup>	SSE treino4	SSE teste4	Média ponderada SSE
24	0,9354	0,8333	<b>0,9150</b>	35,6463	1,88E+03	<b>402,8077</b>	0,9212	0,9024	<b>0,9174</b>	8,3706	197,1717	<b>46,0539</b>
25	0,9278	0,8246	<b>0,9072</b>	746,1546	6,28E+00	<b>598,4807</b>	0,9117	0,9055	<b>0,9105</b>	211,2315	137,4330	<b>196,5019</b>
26	0,6480	0,5490	<b>0,6282</b>			<b>0,0000</b>	5,62E-58	3,71E-58	<b>0,0000</b>			<b>0,0000</b>
27	0,6194	0,5231	<b>0,6002</b>			<b>0,0000</b>	0,1153	8,52E-06	<b>0,0923</b>			<b>0,0000</b>
28	0,9311	0,8186	<b>0,9086</b>	0,0843	1,44E+03	<b>287,0019</b>	0,9149	0,8713	<b>0,9062</b>	0,1335	128,6857	<b>25,7916</b>
29	0,9292	0,8354	<b>0,9105</b>	0,0087	3,47E+02	<b>69,2549</b>	0,9160	0,9014	<b>0,9131</b>	0,0957	180,2786	<b>36,0589</b>
30	0,9315	0,8108	<b>0,9074</b>	1,1703	1,03E+03	<b>206,0182</b>	0,9191	0,8950	<b>0,9143</b>	0,4247	47,1408	<b>9,7489</b>
31	0,9289	0,8157	<b>0,9063</b>	6,8560	46,6965	<b>14,8079</b>	0,9144	0,9301	<b>0,9175</b>	0,7806	318,8826	<b>64,2714</b>
32	0,9301	0,8038	<b>0,9049</b>	0,0499	0,5137	<b>0,1425</b>	0,9146	0,8461	<b>0,9009</b>	0,1372	225,2100	<b>45,0601</b>
33	0,9691	0,0098	<b>0,7776</b>	9,82E-04	3,31E+04	<b>6612,9051</b>	0,9389	0,0259	<b>0,7567</b>	1,0461	4,08E+04	<b>8153,3994</b>

## Apêndice VIII - Valores dos pesos e bias da rede eleita

### Pesos entre camada de entrada e primeira cada intermediária (P1)

P1									
1,0184	0,506	0,502	1,5745	-1,438	-1,5045	-5,2077	-1,5538	0,152	1,7517
-0,4325	0,087	0,187	-2,9708	1,9885	-0,4147	2,2074	-1,8419	0,3235	3,1984
0,5397	1,7434	0,073	0,9192	2,0016	-0,4792	0,9764	3,349	-0,8199	-1,1904
1,3885	0,0882	-1,5408	2,5631	-0,1845	4,7643	-1,8045	-0,4768	-2,3111	-1,8192
-0,6666	0,3263	1,3316	0,2285	-1,8955	1,3136	-2,5958	4,6797	2,9287	-3,0299
-1,9661	-2,6646	1,1016	0,3109	-0,4119	-0,163	-2,9325	-1,8306	-1,8725	1,0635
-0,598	-3,5229	3,0453	-2,1944	0,6439	-0,182	-1,2709	-0,8627	-1,2042	0,5944
0,5827	3,4204	-0,7987	0,413	-2,8713	-1,5929	-2,586	0,0815	-0,2014	1,1342
0,2074	3,6843	-1,1353	-0,0463	1,8402	1,302	1,2978	-1,6328	-2,3735	-3,115
1,0902	-1,5874	-1,0982	1,8743	2,2857	0,051	-0,1947	-0,0644	-2,0531	-1,5626

### Pesos entre primeira cada intermediária e segunda camada intermediária (P2)

P2									
-1,3835	-0,4507	1,4896	0,2905	1,9157	-1,1545	-1,4578	4,6203	-2,8982	2,0277
-1,6505	-1,0297	-3,4832	-1,2845	1,8919	-0,6344	-1,6242	1,0826	-0,3035	1,6797
-3,6809	2,7032	0,8334	-1,3151	-0,4163	0,7865	-2,3741	1,7159	-2,7942	3,0697
0,4871	-1,9041	-4,5910	1,5245	3,2208	1,1832	0,0062	-0,2216	2,1370	-1,9974
-3,8261	-0,2070	-0,8395	1,6151	1,4282	1,1723	-2,3561	1,4075	-1,9332	-1,1966
2,1603	1,1487	-2,5631	-3,4529	-1,7051	-1,0912	-0,0957	0,9949	0,1749	-1,8041
1,1072	-0,2102	2,2899	0,3485	-3,4449	-1,6096	2,8611	-2,1465	-1,4321	2,3410
-1,2432	1,0034	1,9079	-0,6418	1,9713	1,6599	-1,4651	2,2986	-0,3331	-0,6326

P2									
-2,6617	-2,6931	1,4745	0,5652	2,1757	0,0003	-0,2238	-0,2821	0,6602	0,3867
1,4380	-2,3378	0,5009	0,4298	0,4798	1,9774	-1,7536	-0,8645	0,6996	-2,5546

### Pesos entre segunda camada intermediária e camada final (P2)

P3									
3,9206	0,5512	-1,294	-0,1902	-0,0557	-2,0139	-1,2115	0,5806	-2,8435	-3,2709
-3,7465	1,7123	0,1892	-0,6048	-1,8577	2,2961	-0,9	-3,3688	2,5977	1,9872
-2,4747	-2,5564	3,7849	3,2406	-3,4786	0,4982	0,6935	3,4372	0,0909	-0,1937
-2,6264	2,7219	-0,5899	0,1792	-0,4642	-0,4516	2,4097	1,8837	-0,9852	0,4577

### Bias das camadas intermediárias (B1 e B2) e de saída (B3)

B1	B2	B3
-2,2505	3,6542	1,3973
0,0706	0,6664	1,8867
1,3551	2,8893	-3,4106
0,4022	-1,4357	0,5221
1,8162	-0,6524	
0,1332	-0,3655	
0,6949	-2,3508	
-3,0393	-0,6136	
-0,2641	2,1646	
0,0139	1,8163	