

UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA QUÍMICA (PEQ-UFS)

TIAGO HORA ALVES DE LIMA

DESENVOLVIMENTO DE SENSOR VIRTUAL UTILIZANDO REDES NEURAIS
ARTIFICIAIS NA DESTILAÇÃO DE BEBIDAS FERMENTADAS

São Cristóvão (SE)

(2019)

TIAGO HORA ALVES DE LIMA

DESENVOLVIMENTO DE SENSOR VIRTUAL UTILIZANDO REDES NEURAIS
ARTIFICIAIS NA DESTILAÇÃO DE BEBIDAS FERMENTADAS

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Química, como requisito parcial à obtenção do título de Mestre em Engenharia Química.

Orientador: Prof. Dr. Antônio Martins de Oliveira Junior

São Cristóvão (SE)

(2019)

RESUMO

A necessidade de monitorar os processos industriais de forma confiável para garantir a qualidade dos produtos contrasta com o alto custo dos equipamentos disponíveis e a dificuldade de obter informações em tempo real. Em relação às destilarias brasileiras produtoras de aguardente se trata de um desafio ainda maior, visto que a maioria são empresas pequenas, muitas vezes familiares. Neste contexto, sensores virtuais mostram-se uma alternativa viável pela sua capacidade de fornecer informações sobre variáveis de qualidade como a composição a partir de variáveis primárias facilmente mensuráveis, como temperatura e pressão, utilizando modelos matemáticos programados em dispositivos de *hardware*. Com este objetivo, este trabalho desenvolve um sensor virtual para inferir a composição do etanol em um processo de destilação em batelada na produção de bebidas destiladas de frutos tropicais, contornando a dificuldade na execução dos cortes de destilação, realizados manualmente de forma artesanal. O sensor proposto é baseado em uma rede neural artificial do tipo *feedforward multilayer perceptron*, que aplica o algoritmo Levenberg-Marquardt na otimização dos seus parâmetros. Para a sua construção foram utilizados dados de destilações binárias produzidas em laboratório, compostas por etanol e água e de concentração inicial próxima da concentração inicial dos mostos fermentados de frutas. O sensor construído apresentou ótimos resultados, com erro absoluto médio (EAM) de 0,0140 a 0,0311 em fração mássica nos experimentos realizados. Para comprovar a sua eficiência na situação de interesse, o sensor foi testado nas destilações dos mostos fermentados de manga, melancia e jabuticaba, também produzidos em laboratório, obtendo ótima performance, com EAM de 0,0160 a 0,0324. Desta maneira, o sensor virtual proposto mostrou-se capaz de inferir a composição do etanol ao longo do tempo de forma confiável, sendo uma alternativa viável para a realização de um monitoramento eficiente.

PALAVRAS-CHAVE: Bebidas destiladas; sensores virtuais; redes neurais artificiais.

ABSTRACT

The need to monitor industrial processes reliably to ensure products quality contrasts with the high cost of available equipments and the difficulty of obtaining information in real time. In relation to Brazilian distilleries producing spirits, this is an even greater challenge, since most are small, often family-owned enterprises. In this context, virtual sensors are a viable alternative for their ability to provide information about quality variables such as composition from easily measurable primary variables such as temperature and pressure using mathematical models programmed into hardware devices. With this objective, this work develops a soft sensor to infer the composition of ethanol in a batch distillation process in the production of distilled beverages of tropical fruits, overcoming the difficulty in the execution of the distillation cuts, made by a craft way. The proposed sensor is based on an artificial neural network of the feedforward multilayer perceptron type, which applies the Levenberg-Marquardt algorithm in the optimization of its parameters. For its construction were used data from binary distillations produced in the laboratory, composed of ethanol and water and with initial concentration close to the initial concentration of fermented fruit must. The built sensor presented excellent results, with mean absolute error (EAM) of 0,0140 to 0,0311 in mass fraction in the experiments performed. In order to prove its efficiency in the situation of interest, the sensor was tested in the fermented must distillations of mango, watermelon and jabuticaba, also produced in the laboratory, obtaining an excellent performance, with EAM of 0.0160 to 0.0324. In this way, the proposed virtual sensor proved capable of inferring the ethanol composition over time in a reliable way, being a viable alternative for the efficient monitoring.

KEY-WORDS: Distilled beverages; soft sensors; neural networks.

LISTA DE TABELAS

Tabela 1 Parâmetros de qualidade para a aguardente de fruta. Fonte: Brasil - Instrução Normativa MAPA nº 15 de 01 de abril de 2011.....	15
Tabela 2 - Aplicações de sensores virtuais em processos de destilação.....	29
Tabela 3 - Rendimento das destilações do mosto fermentado de melancia.	44
Tabela 4 - Desempenho das melhores Redes Neurais.	50
Tabela 5 - Valores dos pesos e bias da Rede Neural.	53
Tabela 6 - Avaliação de desempenho do sensor virtual - Destilações das misturas sintéticas.	57
Tabela 7 - Análise de desempenho do sensor virtual - Destilações do mosto fermentado de melancia.....	59
Tabela 8 - Análise de desempenho do sensor virtual - Destilações do mosto fermentado de manga.....	61
Tabela 9 - Análise de desempenho do sensor virtual - Destilações do mosto fermentado de jabuticaba.....	63

LISTA DE FIGURAS

Figura 1 - Princípio de trabalho de um sensor virtual. Fonte: Elaborada pelo autor.....	19
Figura 2 - Diagrama de blocos do desenvolvimento de um sensor virtual. Fonte: Elaborada pelo autor.	20
Figura 3 - Neurônio biológico. Fonte: ICMC-USP	31
Figura 4 - Estrutura do neurônio artificial. Fonte: Elaborado pelo autor.	32
Figura 5 - Rede Neural Multicamadas. Fonte: Elaborada pelo autor.	36
Figura 6 - Módulo experimental de destilação. (1) – Refervedor, (2) – Coletor de amostra (presente no refervedor), (3) – Coletor de amostra (produto), (4) – Painel eletrônico, (5) – Módulos de coletas (pratos), (6) – Módulo de topo, (7) – Vacuômetro Bourdon, (8) – Refluxador, (9) – Condensador e (10) – Manta de aquecimento. Fonte: Mota (2017).	42
Figura 7 - Esquema de treinamento da RNA. Fonte: Elaborada pelo autor.	45
Figura 8 - Perfil de temperatura da coluna (a) e de etanol de concentração de etanol no destilado (b) durante o processo - Experimento 1. Concentração inicial: 12,7%(v/v).	47
Figura 9 - Perfil de temperatura da coluna (a) e de concentração de etanol no destilado (b) durante o processo - Experimento 2. Concentração inicial: 13,1 %(v/v).	48
Figura 10 - Perfil de temperatura da coluna (a) e de concentração do etanol no destilado (b) durante o processo - Experimento 3. Concentração inicial: 10,4%(v/v).	48
Figura 11 - Perfil de Temperatura da coluna (a) e de concentração do etanol no destilado (b) ao longo do processo - Experimento 4. Concentração inicial: 12,7 %(v/v).	48
Figura 12 - Perfil de temperatura da coluna (a) e de concentração de etanol no destilado (b) durante o processo- Experimento 5. Concentração inicial: 13,4 %(v/v).	49
Figura 13 - Perfil de temperatura da coluna (a) e de concentração de etanol no destilado durante o processo - Experimento 6. Concentração inicial: 12,4 %(v/v).	49
Figura 14 - Regressão da melhor topologia obtida.....	51
Figura 15 - Curva de treinamento da Rede Neural de melhor desempenho.....	52
Figura 16 - Topologia da Rede Neural de melhor desempenho.	52
Figura 17 - Perfil de concentração do etanol durante o processo de destilação da mistura binária - Experimento 1. Concentração inicial: 12,7%(v/v).	54

Figura 18 - Perfil de concentração do etanol durante o processo de destilação da mistura binária - Experimento 2. Concentração inicial: 13,1 %(v/v).	54
Figura 19 - Perfil de concentração do etanol durante o processo de destilação da mistura binária- Experimento 3. Concentração inicial: 10,4%(v/v).	55
Figura 20 - Perfil de concentração do etanol durante o processo de destilação da mistura binária - Experimento 4. Concentração inicial: 12,7%(v/v).	55
Figura 21 - Perfil de concentração do etanol durante o processo de destilação da mistura binária - Experimento 5. Concentração inicial: 13,4%(v/v).	56
Figura 22 - Perfil de concentração do etanol durante o processo de destilação da mistura binária - Experimento 6. Concentração inicial: 12,4%(v/v).	56
Figura 23 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de melancia - Experimento 1.....	58
Figura 24 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de melancia - Experimento 2.....	58
Figura 25 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de melancia - Experimento 3.....	59
Figura 26 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de manga - Experimento 1.....	60
Figura 27 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de manga - Experimento 2.....	60
Figura 28 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de manga - Experimento 3.....	61
Figura 29 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de jabuticaba - Experimento 1.....	62
Figura 30 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de jabuticaba - Experimento 2.....	62
Figura 31 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de jabuticaba - Experimento 3.....	63
Figura 32 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de jabuticaba - Experimento 4.....	63

SUMÁRIO

1 INTRODUÇÃO	10
2 OBJETIVOS	12
2.1 Objetivo Geral.....	12
2.2 Objetivos Específicos.....	12
3 REVISÃO BIBLIOGRÁFICA	13
3.1 BEBIDAS DESTILADAS.....	13
3.2 SENSORES VIRTUAIS	17
3.2.1 Desenvolvimento de sensores virtuais	19
3.2.2 Sensores virtuais em colunas de destilação	22
3.3 REDES NEURAIS ARTIFICIAIS	29
3.3.1 Neurônio biológico.....	30
3.3.2 Neurônio Artificial	31
3.3.3 Arquitetura de uma Rede Neural.....	33
3.3.5 Aprendizado de uma Rede Neural	34
3.3.6 Redes Perceptron de Múltiplas Camadas	35
3.3.7 Algoritmo de retropropagação do erro (<i>backpropagation</i>).....	35

3.3.8 Algoritmo Levenberg-Marquardt.....	38
4 METODOLOGIA.....	41
4.1 ETAPA EXPERIMENTAL	41
4.1.1 Materiais utilizados	41
4.1.2 Destilações das misturas sintéticas.....	42
4.1.3 Produção do mosto fermentado.....	43
4.1.4 Destilação do mosto fermentado e produção da aguardente de fruta.....	43
4.2 ETAPA COMPUTACIONAL	44
5 RESULTADOS E DISCUSSÕES.....	47
5.1 FORMAÇÃO DO CONJUNTO DE DADOS	47
5.3 TESTES COM AS DESTILAÇÕES DE MOSTOS FERMENTADOS	57
6 CONCLUSÃO.....	65
REFERÊNCIAS.....	66

1 INTRODUÇÃO

Bebidas destiladas são populares no mundo inteiro por conta do seu sabor e aroma distinto, sendo muitas vezes produzidas através de frutos tradicionalmente regionais. Destilados de frutas são misturas complexas que consistem em etanol, água e centenas de componentes voláteis chamados associados ou congêneres, cuja concentração determina as características sensoriais do produto, como aroma e sabor (OSORIO *et al.*, 2005; LUNA *et al.*, 2018).

A maior parte desses componentes é formada durante a fermentação e possui diferentes graus de volatilidade, o que permite a sua separação a partir da destilação. Tal separação, essencial na determinação da qualidade do produto final, é realizada através de cortes, cabeça, coração e cauda, que variam de acordo com o teor de etanol (ALCARDE *et al.*, 2010). Entretanto, muitas vezes os cortes são realizados manualmente através de alcoometria ou sensorialmente, o que impossibilita a padronização do processo e ocasiona em perdas de qualidade e produtividade (CLAUS & BERGLUND, 2009).

Para realizar um monitoramento e controle eficiente do processo de destilação, faz-se necessário obter informações sobre a composição dos produtos em tempo real, a fim de diminuir o tempo de resposta dos controladores e aumentar a eficiência na tomada de decisões. Usualmente, as indústrias utilizam analisadores de linha como hidrômetros e cromatógrafos, mas tal alternativa apresenta desvantagens como atrasos nas informações, o que impossibilita o monitoramento dos processos em tempo real, além de demandar altos investimentos na aquisição de equipamentos e necessitar de manutenção constante (SHAO & TIAN, 2015).

Diante da necessidade de monitorar os processos industriais para obter ganhos em qualidade e diminuir os custos de produção, os sensores virtuais são uma alternativa viável na realização do monitoramento online de forma barata e eficiente.

Sensores virtuais são modelos matemáticos capazes de inferir variáveis de processo difíceis de serem mensuradas como a composição através de variáveis facilmente mensuráveis como temperatura, pressão e fluxo (TORGASHOV & ZMEU, 2015). Tais sensores possuem a vantagem de serem implementados em dispositivos de hardware e viabilizam o monitoramento de variáveis importantes quando por algum motivo não existem dispositivos disponíveis para realizar tal tarefa com a precisão ou frequência desejadas (FORTUNA *et al.*, 2007).

Entre as principais abordagens na modelagem de sensores virtuais, fenomenológica e empírica (KADLEC *et al.*, 2009), percebe-se a predominância da aplicação de sensores virtuais empíricos. Isto acontece devido à dificuldade em realizar modelagens rigorosas por conta da complexidade dos processos industriais, da alta demanda computacional desses modelos e da grande quantidade de dados disponíveis nas indústrias (ABDULLAH *et al.*, 2007; GRBIC *et al.*, 2013). Neste contexto, o presente trabalho desenvolve um sensor virtual empírico baseado em Redes Neurais artificiais para inferir a composição de etanol em um processo de destilação em batelada na produção de destilados de frutas.

Deste modo, uma Rede Neural do tipo *feedforward multilayer perceptron* foi desenvolvida para compor o sensor. Para a sua calibração, destilações de misturas binárias etanol-água com concentrações iniciais de etanol próximas das concentrações dos mostos fermentados foram realizadas. Para verificar a sua capacidade de inferência na situação de interesse, o sensor foi testado com as destilações dos mostos fermentados de manga, melancia e jabuticaba, também produzidos em laboratório nas mesmas condições operacionais que as misturas sintéticas.

O trabalho está organizado em seis seções. A primeira diz respeito a esta introdução e na segunda constam os objetivos geral e específicos. Na terceira seção está o referencial teórico utilizado neste trabalho, que trata de bebidas destiladas, sensores virtuais e redes neurais artificiais. As metodologias experimental e computacional aplicadas no desenvolvimento do sensor estão descritas na quarta seção. Na quinta seção são expostos os resultados obtidos e as respectivas discussões e na sexta as conclusões e propostas para trabalhos futuros.

2 OBJETIVOS

2.1 Objetivo Geral

Desenvolver e validar um sensor virtual baseado Redes Neurais Artificiais para inferir a composição do etanol em processos de destilação de fermentados de frutas em batelada.

2.2 Objetivos Específicos

- Desenvolver um sensor capaz de prever a composição de etanol em processos de destilação de mostos fermentados de frutas;
- Treinar e validar uma Rede Neural Artificial utilizando dados de uma mistura binária para ser utilizada como sensor virtual;
- Analisar o desempenho da Rede Neural Artificial desenvolvida na inferência da composição do etanol ao longo da destilação de mostos fermentados.

3 REVISÃO BIBLIOGRÁFICA

Neste capítulo está a revisão bibliográfica com os assuntos relevantes para o tema da dissertação. A primeira parte trata das bebidas destiladas, a segunda dos sensores virtuais e a terceira das Redes Neurais Artificiais.

3.1 BEBIDAS DESTILADAS

Destilados de frutas são muito populares no mundo inteiro e frequentemente são derivados de frutos tradicionalmente regionais. Existem diversos exemplos de bebidas tradicionais como a Grappa, destilado derivado de uva popular na Itália (GALANO *et al.*, 2015), o Pisco, destilado de uva Muscat produzido no Peru e Chile (OSORIO *et al.*, 2005), a Sljivovica, destilado do suco fermentado de ameixas comum no leste europeu (PECIC *et al.*, 2012), etc.

No Brasil, terceiro maior produtor de frutos do mundo (OECD, 2018), também são encontrados destilados derivados de frutos como a manga (CALIARI *et al.*, 2012), banana (SILVA *et al.*, 2009), jabuticaba (ASQUIERI *et al.*, 2009), etc. Neste contexto, além de ser um produto tradicional, os destilados de fruta também representam uma alternativa econômica sustentável ao desperdício dos alimentos que não são absorvidos pelo mercado, movimentando a atividade econômica e agregando valor às bebidas regionais (ASQUIERI *et al.*, 2009, PARENTE, 2014).

Segundo a legislação brasileira, aguardente de fruta é a bebida com graduação alcoólica de trinta e seis a cinquenta e quatro por cento em volume, a vinte graus Celsius, obtida de destilado alcoólico simples de fruta ou pela destilação de mosto fermentado de fruta (BRASIL, 2009). Compostas em sua maior fração por etanol e água, cerca de 96% em massa (OSORIO *et al.*, 2005), as bebidas destiladas, assim como as aguardentes de fruta, possuem uma série de componentes voláteis em menor concentração, chamados de associados ou congêneres, cuja concentração determina as características sensoriais, como aroma e sabor, da bebida destilada. A concentração desses componentes no produto final depende de diversos fatores, tais como a matéria prima utilizada, o processo de fermentação, a destilação e a técnica de envelhecimento aplicada após a destilação, que pode ser realizada em alambiques ou em colunas, em batelada ou de forma contínua (GARCIA-LLOBODANIN *et al.*, 2007; SACHER *et al.*, 2013).

Segundo Spaho *et al.* (2013), apesar da importância das etapas anteriores como a seleção e tratamento adequado da matéria-prima e o processo de fermentação na qualidade do produto final, falhas nesses processos podem ser corrigidas em etapas posteriores, como a destilação e o envelhecimento, o que é muito difícil de realizar caso a execução do processo de destilação seja inapropriada, resultando em produtos com características desagradáveis.

A maior parte dos componentes voláteis é formada durante a fermentação e durante a destilação os componentes desejáveis, que agregam as características sensoriais de interesse à bebida, são separados dos componentes indesejáveis. Segundo Batista e Meireles (2011), os principais componentes formados durante a fermentação são álcoois, (metanol, etanol, propanol e álcool isoamílico), ácidos orgânicos (ácido acético), compostos carbonila (acetaldeído) e ésteres (acetato de etila).

Conforme o estudo de Leauté (1990), os vinhos contêm centenas de componentes voláteis que destilam de acordo com três critérios: ponto de ebulição, a sua interação com a água e o etanol e consequentemente a variação do teor de etanol e de água durante o processo. Os componentes solúveis ou parcialmente solúveis em etanol são destilados quando o vapor que flui ao longo da coluna é rico em etanol, os componentes de maior afinidade com a água quando o vapor é mais rico em água, os componentes solúveis tanto em água quanto em etanol são destilados durante todo o processo e os componentes insolúveis são carregados pelo fluxo de vapor.

Diante dessas características, a destilação de mostos fermentados é separada em três cortes distintos de acordo com o seu teor de etanol: cabeça, coração e cauda. É importante ressaltar que a composição dos cortes varia de acordo com a matéria-prima utilizada. Segundo Spaho (2017), os mostos fermentados de diferentes frutas possuem algumas características distintas e os métodos de destilação precisam ser ajustados neste sentido.

A fração cabeça contém os componentes mais voláteis, que são separados no início da destilação. Possuem baixo ponto de ebulição e são solúveis em etanol. O coração, produto a ser comercializado, possui componentes de moderada a alta volatilidade, solúveis em etanol e solúveis ou parcialmente solúveis em água. A cabeça contém compostos muito solúveis em água e de alto ponto de ebulição (ALCARDE *et al.*, 2010).

A Tabela 1 define a concentração dos componentes menores permitidos no coração de acordo com a legislação brasileira.

Tabela 1 Parâmetros de qualidade para a aguardente de fruta.

Parâmetro de qualidade	Legislação brasileira
Grau alcoólico real a 20°C (v/v)	36-54
Acidez volátil em ácido acético	0-150
Aldeídos em aldeído acético	máx 30
Ésteres em acetato de etila	máx 250
Álcool metílico	<20
Álcool sec-butanol	0-10
Álcool n-butílico	0-3
Álcoois superiores	máx 360
Furfural	máx 5
Coeficiente de congêneres	200-650
Teor de açúcar (g/L)	< 6

Fonte: Brasil - Instrução Normativa MAPA nº 15 de 01 de abril de 2011.

Devido a importância do processo de destilação na determinação das características sensoriais da bebida, diversos estudos em destilação de bebidas alcóolicas vêm sendo realizados. Hernandez-Gomes *et al.* (2003) estudaram a influência do tipo de destilador e do tratamento da matéria-prima na qualidade do produto final. Em uma abordagem semelhante Arrieta-Garay *et al.* (2014) investigaram o perfil aromático da Grappa de acordo com o tipo de destilador, alambique de cobre ou coluna, e a levedura utilizada no processo de fermentação. Madrera *et al.* (2003) estudaram como a destilação, a madeira e o tempo aplicados no envelhecimento influenciam o buquê de destilados de cidra.

Em outra abordagem, Spaho *et al.* (2013) verificou se a espécie de ameixa ou a fração de etanol nos cortes tem maior influência sobre o teor de componentes voláteis na produção de bebidas utilizando dupla destilação. Também por meio de dupla destilação, Alcarde *et al.* (2010) analisaram a cinética de volatilidade dos componentes associados no processo de produção de cachaça. Os experimentos mostraram que os aldeídos, ésteres e metanol se concentram na fração cabeça, o ácido acético na fração cauda e que os álcoois superiores se concentram nas frações alcóolicas entre 35 e 55%(v/v).

No estudo de Bruno *et al.* (2007), foi analisado o teor de carbamato de etila em diversas destilarias do Rio de Janeiro. O carbamato de etila é um componente tóxico e cancerígeno permitido em pequenas quantidades em destilados como a aguardente. As análises foram realizadas dividindo as amostras de acordo com o método de destilação utilizado, como alambiques de aço, cobre e colunas de destilação contínuas. Já Lima *et al.* (2012) pesquisaram a influência do tempo de destilação no teor de carbamato de etila na bebida destilada, constatando que tempos maiores de destilação acarretam em menor formação da substância. Li

et al. (2015), também investigaram a formação de carbamato de etila. Em abordagem semelhante à de Bruno *et al.* (2007), os autores analisaram amostras de diversas destilarias chinesas quanto ao teor de carbamato de etila e os seus métodos de produção. Foi constatado que a natureza da matéria prima, o tempo de fermentação, armazenamento e a sua temperatura são os parâmetros de maior influência na sua formação.

Além de abordagens heurísticas com foco nas características sensoriais do produto decorrentes do método de destilação utilizado, também encontram-se na literatura pesquisas referentes à modelagem e simulação de processos de destilação de bebidas alcoólicas, cujo objetivo é entender o comportamento de determinados componentes da mistura ao longo do processo de destilação a fim de viabilizar a elaboração de estratégias de operação que maximizem o teor de substâncias favoráveis à sua qualidade e minimizem substâncias tóxicas e/ou prejudiciais ao aroma e sabor do produto.

Tal finalidade foi abordada nos trabalhos de Scavanini *et al.* (2012), que estudaram o comportamento do etanol e diversos componentes associados em um processo de destilação de cachaça em alambique e Sacher *et al.* (2013), que modelaram o comportamento do etanol e diversos congêneres na destilação de um mosto fermentado de pêra. Esteban-Decloux *et al.* (2014) estudaram o comportamento de etanol e congêneres em dois casos, um processo de destilação contínua na produção de álcool neutro e um processo de destilação em batelada na produção de um destilado de laranja. Decloux e Coustel (2005), produção de álcool neutro em colunas de destilação, Claus e Berglund (2009), produção de destilado de maçã em batelada, Batista e Meirelles (2011), produção de cachaça em destilação contínua e Puentes *et al.* (2018), produção de armanhaque, também desenvolveram modelos com esse objetivo.

Estratégias de controle foram abordadas por Osorio *et al.* (2005) e Luna *et al.* (2018). Em seu trabalho, Osorio *et al.* (2005) resolveram um problema de controle com o objetivo de maximizar a concentração de terpenos, substâncias responsáveis pelo aroma característico do Pisco, e minimizar o teor de metanol e ácidos orgânicos, prejudiciais à qualidade do produto, através de um método de otimização sequencial. O problema consiste em encontrar uma trajetória ótima de controle para o condensador e o refeedor. Desta maneira foi possível chegar a um produto com teor de terpenos mais próximo do indicado por enologistas que as bebidas comerciais, além de baixas concentrações de componentes indesejáveis.

Para minimizar o teor de metanol no produto final, Luna *et al.* (2018) propuseram a implementação de um sistema de controle de processos que envolve o desenvolvimento de um

modelo dinâmico. Para tal, os autores utilizaram um modelo fenomenológico presente na literatura auxiliado por equações que descrevem a perda de energia para o ambiente e o equilíbrio líquido/vapor da mistura, validado por experimentos em um alambique, obtendo redução significativa do teor de metanol.

Como frisado por Puentes *et al.* (2018), apesar da implementação de simuladores na indústria de alimentos ser relativamente escassa, muito vem sendo publicado nos últimos anos. Entretanto, para que as estratégias de controle desenvolvidas sejam implementadas com sucesso, um monitoramento eficiente da composição do destilado se faz necessário. De acordo Spaho (2017) os métodos mais utilizados no monitoramento da destilação de mostos ainda são a análise sensorial e a alcoometria. Tais procedimentos impossibilitam a repetibilidade do processo ocasionando em perdas de qualidade e produção devido à falta de padronização. Segundo Bortoletto e Alcarde (2015), uma das razões que dificulta a exportação da cachaça, destilado mais consumido no Brasil, é a baixa qualidade e a falta de padronização do produto. Em seu estudo, Coutinho *et al.* (2012) ressaltam que em várias regiões do Brasil a aguardente é produzida de maneira empírica por meio de métodos arcaicos e o conhecimento ainda é passado de pai para filho em empresas familiares.

Outro problema na realização do monitoramento é a falta de instrumentos de baixo custo para a realização desta tarefa. A análise da composição pode ser realizada por hidrômetros ou cromatógrafos, equipamentos que necessitam de investimento na sua aquisição e manutenção. Tendo em vista o cenário das destilarias brasileiras, no qual 99% é formado por produtores pequenos (DCI, 2017), esta recomendação pode se mostrar inviável. Neste contexto, para realizar a tarefa de monitoramento da composição no processo de destilação de bebidas alcóolicas de forma barata e confiável, sensores virtuais são uma alternativa viável.

3.2 SENSORES VIRTUAIS

As plantas industriais são equipadas com uma grande variedade de sensores cujo objetivo é fornecer informações sobre o seu estado para os sistemas de monitoramento e controle de processos. As áreas de aplicação desses sensores são muito diversas, como química, petroquímica, mineração, siderurgia, farmacêutica, etc. Logo, os sensores funcionam de acordo com a variável que deseja ser medida e com as condições de operação do processo (LOTUFO e GARCIA, 2008).

Segundo Sliskovic *et al.* (2011), o objetivo dos sistemas de monitoramento e controle dos processos industriais é garantir a qualidade desejada dos produtos, reduzir custos de operação, garantir o funcionamento seguro da planta e a eficiência desses sistemas depende da qualidade das medidas fornecidas pelos sensores. Entretanto, algumas variáveis, geralmente variáveis que carregam informações sobre a qualidade do produto final, por algum motivo não estão disponíveis em tempo real, seja pela falta, manutenção ou não existência do sensor, o que prejudica o monitoramento do processo e impossibilita o seu controle automático.

Em diversas situações, as informações sobre as variáveis de qualidade são obtidas por meio de analisadores de linha instalados na planta, mas tais dispositivos possuem um alto custo e podem extrapolar o orçamento das indústrias, além de necessitar de manutenção constante. Instrumentos como hidrómetros e cromatógrafos apresentam atrasos nas medidas, baixa precisão e fornecem informações distorcidas conforme mudam as condições de operação (KANO *et al.*, 2000; OSORIO *et al.*, 2008). Uma alternativa aos analisadores de linha são as análises laboratoriais de amostras obtidas na planta conduzidas *off-line*, mas além de fornecer resultados em uma frequência baixa, demandam investimento em recursos humanos, equipamentos caros e nem sempre são aplicáveis em escala industrial (GRBIC *et al.*, 2013; TOGASHOV e ZMEU, 2015; SHAO e TIAN, 2015).

Nos últimos anos, pesquisadores vêm utilizando a grande quantidade de dados gerados e armazenados pelas plantas industriais na construção de modelos preditivos que fornecem informações sobre variáveis difíceis de mensurar em tempo real, estes modelos são chamados sensores virtuais. De acordo com Togashov e Zmeu (2015), sensores virtuais (do inglês “*soft sensor*”), analisadores virtuais ou sensores inferenciais, são modelos matemáticos capazes de estimar variáveis chave do processo difíceis de serem mensuradas, geralmente variáveis de qualidade, utilizando variáveis secundárias facilmente mensuráveis como fluxo, temperatura, pressão, etc.

Doraiswami e Cheded (2014) destacam as seguintes vantagens no uso de sensores virtuais:

- São uma alternativa barata aos sensores de *hardware*;
- São confiáveis, não necessitam de manutenção e não representam perigo físico, pois são baseados em *softwares*;
- Agregam qualidade aos produtos, pois podem medir variáveis de qualidade que geralmente são de difícil mensuração;

- São úteis na fusão de dados, possuindo a capacidade de fornecer informações sobre uma variável a partir de dados de diversas variáveis;
- As suas principais aplicações são a predição *online*, detecção de falhas e monitoramento do processo, podendo ser implementado a um controlador estimando variáveis da planta difíceis de serem medidas.

O termo *soft sensor* deriva das palavras *software*, que se refere aos modelos utilizados advindos de programação computacional, e *sensor*, referente aos sensores de *hardware* instalados nas plantas industriais que fornecem as informações utilizadas na predição pelo sensor virtual. Ou seja, a ideia básica do sensor virtual é a relação entre sensor e modelo (KADLEC *et al.*, 2009; LOTUFO e GARCIA, 2008).

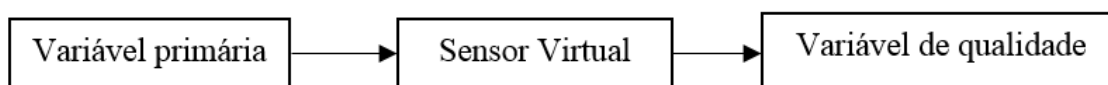


Figura 1 - Princípio de trabalho de um sensor virtual. Fonte: Elaborada pelo autor.

De acordo com Lotufo e Garcia (2008), os sensores virtuais são resultado da intersecção da tecnologia de sensores inteligentes com as técnicas de modelagem e identificação de sistemas. Eles surgem como alternativa para resolver problemas como imprecisão, atraso, preço, disponibilidade e confiabilidade dos sensores de *hardware*, além das dificuldades impostas pelo ambiente e distância do ponto de medição.

3.2.1 Desenvolvimento de sensores virtuais

Segundo Souza *et al* (2016), o desenvolvimento de sensores virtuais segue o ciclo clássico da identificação de sistemas. As etapas ilustradas pela Figura 2 são abordadas nos trabalhos de Fortuna *et al.* (2007), Kadlec *et al.* (2009) e Souza *et al.* (2016) e serão descritas a seguir.

- Coleta e filtragem de dados

O primeiro passo do desenvolvimento de sensores virtuais é a coleta dos dados necessários para a sua modelagem. No primeiro contato, é importante ter uma visão geral sobre a estrutura dos dados disponíveis, pois é a partir do comportamento dos dados que é feita a escolha do modelo. Ao trabalhar com dados reais de processos industriais, é comum lidar com alguns problemas como a presença de *outliers*, dados faltantes, taxa de amostragem não sincronizada, etc. Tais ocorrências podem prejudicar o desempenho do sensor virtual caso não

sejam identificadas e tratadas com a técnica adequada. Algumas das técnicas mais usadas são a regra três sigma (tratamento de *outliers*), imputação de dados (dados faltantes) e resposta de impulso finita (taxa de amostragem variável) (PANI e MOHANTA, 2011; SOUZA *et al*, 2016). O tratamento dos dados deve seguir até que o projetista entenda que o conjunto de dados disponível tem qualidade suficiente para treinar e desenvolver o modelo (ABONYI *et al*, 2014).

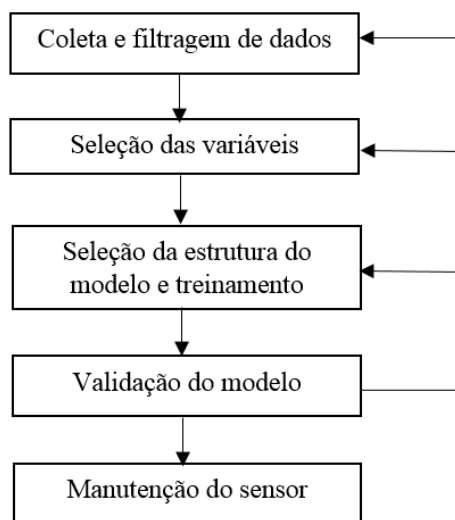


Figura 2 - Diagrama de blocos do desenvolvimento de um sensor virtual. Fonte: Elaborada pelo autor.

- Seleção das variáveis

Os conjuntos de dados das plantas industriais possuem informações sobre todas as variáveis do processo. De acordo com Pani e Mohanta (2011), é necessário identificar quais variáveis primárias influenciam na variável desejada, a fim de evitar um número excessivo de parâmetros no modelo, além de problemas como *overfitting*, quando o modelo apresenta bom desempenho no treinamento e pobre quando apresentado a um novo conjunto de dados, e otimizar o tempo necessário para o seu treinamento. Técnicas de estatística multivariada como PCA (Análise de Componentes Principais) e PLS (Mínimos Quadrados Parciais) são muito utilizadas na identificação de variáveis descartáveis quando apenas o conhecimento sobre o processo não é capaz de determinar a correlação com a variável de interesse.

- Seleção da estrutura do modelo e treinamento

Nesta etapa, deve-se analisar qual abordagem é mais adequada no desenvolvimento do modelo que deverá descrever o processo. Na literatura, os sensores virtuais são classificados de acordo com a natureza do modelo utilizado. Sendo assim, são divididos em dois grandes grupos, os sensores virtuais orientados por modelos e os sensores virtuais orientados por dados. Outras

nomenclaturas também podem ser encontradas, como sensores virtuais paramétricos e não paramétricos, caixa branca e caixa preta, fenomenológicos e empíricos, etc (KADLEC, 2014).

Os sensores virtuais orientados por modelos são baseados em modelos fenomenológicos, ou seja, modelos compostos por equações que descrevem o comportamento do sistema, tais como balanços de massa e energia, equações de equilíbrio, equações de conservação de massa, cinética de reações, etc (KADLEC *et al.*, 2009).

Devido essas características, tais modelos necessitam de um conhecimento *a priori* sobre o comportamento dos sistemas. Tal pré-requisito traz algumas limitações ao uso de modelos fenomenológicos em sensores virtuais, pois processos industriais são muito complexos e uma abordagem teórica muito rigorosa pode se tornar impraticável (GRBIC *et al.*, 2013). Fortuna *et al.* (2005) ressaltam que dependendo da complexidade do fenômeno envolvido pode haver um número excessivo de parâmetros a serem determinados, o que dificulta ainda mais uma modelagem analítica eficiente.

Outro empecilho para a aplicação de modelos fenomenológicos é o seu alto custo computacional. Segundo Abdulah *et al.* (2007), os modelos de primeiro princípio podem apresentar um alto número de equações, tornando-se complexos demais para serem usados nos projetos de controle baseados em modelos não lineares e aumentando consideravelmente o custo computacional do controlador.

Em contraste com sensores de caixa branca, sensores virtuais orientados por dados, ou modelos de caixa preta, baseiam-se na relação de entrada e saída dos dados do processo, dispensando a necessidade do conhecimento *a priori* sobre as suas leis fundamentais, ou seja, o modelo gerado é uma estrutura que transcreve o comportamento do sistema sem fornecer informações funcionais sobre ele ou os seus componentes (LUCAS, 2012). Por conta dessas vantagens e das numerosas instalações de sistemas de controle que armazenam uma grande quantidade de dados sobre os processos industriais, métodos empíricos vem sendo muito utilizados no desenvolvimento de sensores virtuais (SHI e XIONG, 2018).

Nesta abordagem são utilizadas técnicas matemáticas para extrair informações dos dados e descrever toda a dinâmica do processo. Para isto, o tipo de modelo a ser utilizado deve ser analisado, se linear ou não linear. Segundo Souza *et al.* (2016), é recomendável utilizar sempre um modelo linear devido a sua modelagem mais simples. A escolha de modelos não lineares deve ser feita apenas no caso dos modelos lineares não serem aplicáveis. Entretanto, processos industriais geralmente apresentam comportamento não linear.

As principais técnicas utilizadas na modelagem de sensores virtuais empíricos são Análise de Componentes Principais e Mínimos Quadrados Parciais, utilizadas na construção de modelos lineares, Redes Neurais Artificiais (RNA), Sistemas Neuro-Difusos (*Fuzzy Systems*) e Máquinas de Vetores de Suporte (SVM), aplicadas no desenvolvimento de modelos não lineares (KADLEC *et al.*, 2009; SHAO e TIAN, 2015; SOUZA *et al.*, 2016).

Em muitos processos industriais, existe relativo conhecimento sobre o processo, contudo, este é incompleto. A modelagem de caixa cinza, também conhecida como modelagem híbrida, situa-se entre os extremos da modelagem puramente empírica e a modelagem fenomenológica, combinando ambas as metodologias. Segundo James *et al.* (2000) o fundamento da modelagem híbrida é mesclar o conhecimento do processo com as vantagens do mapeamento de dados de entrada e saída. De acordo com Henson (1998) existem duas maneiras típicas de construir modelos híbridos, utilizar modelos empíricos para estimar funções desconhecidas dos modelos fundamentais, como taxas de reação em um modelo de reator químico e aplicar modelos fenomenológicos para capturar características básicas do processo enquanto descreve-se o resíduo entre a planta e o modelo a partir de modelos empíricos.

Em todo caso, a escolha do modelo, seja ele fenomenológico, empírico ou híbrido, depende intrinsecamente de fatores como a disponibilidade do conhecimento necessário para a modelagem do sistema, a qualidade dos dados históricos do processo e a experiência do projetista.

3.2.2 Sensores virtuais em colunas de destilação

Manter a composição do destilado em uma faixa para fins de atender as exigências do mercado agregando valor aos produtos é o principal objetivo de controlar um processo de destilação. No entanto, usualmente as informações relativas à composição de destilado são difíceis de serem obtidas em tempo real, fazendo-se recorrer ao uso de analisadores de linha e análises *off-line*. Nos últimos anos, muitos pesquisadores têm desenvolvido sensores virtuais para estimar a composição de produtos em colunas de destilação e diferentes metodologias têm sido aplicadas.

Oisiovici e Cruz (2000) modelaram sensores utilizando Filtro de Kalman a fim de estimar a composição dos produtos de uma coluna de destilação em batelada. Um modelo não linear com número reduzido de equações foi utilizado para compor o algoritmo do Filtro de Kalman. Valores de temperaturas em diversos pontos ao longo da coluna foram utilizados como

entrada no modelo, assim como a influência do número de entradas e a sua frequência de amostragem foram analisadas quanto ao resultado final. Foi observado um número ótimo de entradas e taxa de amostragem que leva a melhores precisão e convergência dos estimadores. Os sensores se mostraram eficientes na representação da dinâmica do sistema.

Zamprogna *et al.* (2002) desenvolveram modelos utilizando mínimos quadrados parciais (PLS) para inferir as composições de topo e de fundo de uma coluna de destilação em batelada a partir de medidas de temperatura. Mais conhecidos por ser uma técnica de aproximação de funções lineares, nesta abordagem uma função de natureza não linear, que pode ser um regressor polinomial de segunda ordem ou uma rede neural artificial, é aplicada à variável independente na transformação vetorial do PLS para representar a relação não linear existente entre as variáveis de entrada e saída. Os dados utilizados na construção do sensor foram gerados a partir de um simulador com modelos determinísticos variando as condições de operação. Os resultados indicaram que o sensor virtual baseado no PLS não linear se mostrou muito mais preciso que o PLS tradicional.

Para estimar o teor de pentanos no topo e a composição de butanos no fundo de uma coluna debutanizadora, contornando os problemas com atrasos no monitoramento e controle do processo devido ao uso de cromatografia gasosa utilizada anteriormente na avaliação da qualidade dos produtos, Fortuna *et al.* (2005) desenvolveram sensores virtuais. Neste sentido, foram construídos modelos utilizando Redes Neurais Artificiais a partir de dados históricos dos últimos três anos de operação da planta industrial. Os resultados mostraram que os sensores fornecem respostas rápidas e são uma alternativa barata para a estimação das variáveis de qualidade em tempo real, capazes de substituir de forma eficiente os cromatógrafos.

Zhang *et al.* (2004) propuseram um sensor virtual para resolver o problema de monitoramento do teor de água no prato de fundo de uma coluna de destilação azeotrópica. Foram utilizados dados do processo obtidos através do sistema de controle digital da planta e o modelo foi construído aplicando análise regressiva. Com o objetivo de melhorar a sua performance, o sensor foi sendo corrigido por meio de novos dados obtidos em análises laboratoriais em um determinado intervalo. Após meses de uso concluiu-se que o sensor é viável e disponível para uso por longos períodos de tempo.

Em situação semelhante à proposta pelo presente trabalho, Osorio *et al.* (2008) trabalharam com Redes Neurais na construção de um sensor virtual para estimar a composição de etanol em um processo de destilação em batelada na produção de Pisco. Os autores utilizaram

redes neurais artificiais do tipo diretas e recorrentes no desenvolvimento dos modelos. Para treinar e validar as redes neurais, dados reais de destilações em escala industrial e em laboratório foram utilizados, que passaram por pré-processamento a fim de reduzir os efeitos dos ruídos e dos atrasos nas medições. O pré-processamento dos dados melhorou a precisão dos sensores, que mostraram-se uma alternativa barata e confiável na estimação da composição do etanol para as destilarias de vinho.

Smets *et al.* (2010) testaram o modelo de Friedman (1995) na construção de um sensor virtual híbrido para inferir a composição dos produtos em dois processos de destilação multicomponentes. O modelo é constituído de quatro equações, limitando o sensor a ser utilizado em processos onde deseja-se determinar a composição de no máximo quatro componentes. A modelagem é híbrida pois os parâmetros do modelo são determinados pelos dados do processo. A abordagem proposta apresentou bons resultados e foi comparada a um sensor puramente empírico baseado em mínimos quadrados parciais (PLS). Quando apresentado a conjuntos de dados menores, o sensor híbrido tem menor queda de performance que o sensor PLS. Porém, foram observadas dificuldades no ajuste em situações de baixas concentrações, desvantagem não apresentada pelo sensor PLS.

Mohler *et al.* (2011) desenvolveram um sensor virtual para determinar a qualidade dos produtos de uma coluna de destilação de óleo cru (CDU). Tal informação é fornecida pela variável D95 (ponto final da destilação), que informa a quantidade de água e betuminosos em uma destilação de petróleo. Métodos de regressão como PLS e PCA foram utilizados na definição das variáveis de entrada e os dados, coletados por analisadores de linha na planta e análises laboratoriais, passaram por pré-processamento. Diferentes modelos dinâmicos lineares e não lineares foram desenvolvidos utilizando vários métodos de identificação. Os modelos baseados no método Box-Jenkins, método do erro de saída e no modelo de Hammerstein–Wiener apresentaram os melhores resultados.

Em outra abordagem utilizando Redes Neurais, Rani *et al.* (2011) desenvolveram sensores virtuais para estimar a composição de produtos em dois processos de destilação, multicomponentes e reativa, em uma refinaria. Os autores utilizaram Redes Neurais do tipo ADALINE e redes neurais multicamadas com algoritmo Levenberg-Marquardt no seu desenvolvimento, que foi fundamentado em uma base de dados proveniente de simuladores robustos. Os sensores ADALINE apresentaram um melhor desempenho, com melhor precisão e menor custo computacional.

Em outro estudo relacionado a uma unidade de destilação de óleo cru, Rogina *et al.* (2011) empregaram sensores virtuais para inferir a qualidade da nafta leve. A qualidade da nafta é avaliada pela pressão de vapor Reid (RVP), variável inferida pelos sensores virtuais com a finalidade de diminuir os atrasos provocados pelas análises laboratoriais, que determinam a RVP na planta real. Os sensores foram construídos utilizando modelos baseados em Redes Neurais Artificiais do tipo perceptron de múltiplas camadas (MLP) e funções de base radial (RBF), treinadas a partir de dados da planta industrial coletados durante um ano. O sensor virtual baseado na rede MLP apresentou melhor desempenho que o RBF, e a sua aplicação mostrou-se uma alternativa viável para contornar os atrasos provocados pelas análises laboratoriais.

Ramli *et al.* (2014) utilizaram equações baseadas nos modelos de redes neurais artificiais na implementação de sensores virtuais para estimar a composição de n-butano no topo e no fundo de uma coluna debutanizadora. Os autores comentam que tais equações são mais robustas e confiáveis que os modelos de caixa pretas convencionais utilizados na aplicação de redes neurais artificiais. Algumas informações sobre variáveis de entrada e de saída não estavam disponíveis e foram geradas através do simulador HYSYS, a outra parte dos dados utilizadas foram da própria planta industrial. Foram utilizados modelos baseados em redes neurais auto regressivas de entradas exógenas (NARX) com algoritmo de treinamento Levenberg-Marquardt, mínimos quadrados parciais (PLS) e análise regressiva (RA), tendo o sensor NARX apresentado desempenho superior aos demais.

Com o objetivo de inferir a composição dos produtos de um processo de destilação em batelada Yelgatte *et al.* (2015) utilizaram redes neurais artificiais do tipo funções de base radial no desenvolvimento de um sensor virtual. No treinamento destas Redes Neurais, os autores aplicaram o método dos mínimos quadrados ortogonais. O modelo desenvolvido teve como entradas temperaturas de três pratos da coluna e a partir delas fornecia os valores das frações molares dos componentes leves e intermediários no destilado e do componente mais pesado no refeedor. Os dados foram gerados a partir de modelos fenomenológicos que continham sistemas de equações algébricas e diferenciais. As relações de equilíbrio líquido-vapor foram consideradas seguindo o modelo UNIQUAC. Os resultados indicaram que o sensor descreve o processo de destilação com precisão.

Togashov e Zmeu (2015) modelaram um sensor virtual para prever a concentração de isopentanos no topo de uma coluna de destilação contínua de uma unidade de craqueamento catalítico fluido. Os autores aplicaram modelagem de equações estruturais no desenvolvimento

do sensor para lidar com a incerteza em relação à estrutura do modelo e as não linearidades do processo. Na identificação desta estrutura, foram utilizados o algoritmo ACE, um método estatístico usualmente aplicado na resolução de problemas de regressão não lineares, e o coeficiente de identificabilidade, que determina se o modelo é identificável ou não a partir do número de parâmetros e variáveis e se a variável desejada pode ou não ser estimada. Foram utilizados dados gerados por simuladores rigorosos e dados históricos da planta industrial na construção do modelo. Os resultados obtidos comprovaram a eficiência do sensor e a sua viabilidade no uso em escala industrial.

Em seu estudo, Patil e Ghate (2015) propõem um sensor virtual para inferir a composição dos produtos de um processo de destilação multicomponentes. Para isto, redes neurais do tipo *feedforward* com algoritmo Levenberg-Marquardt e redes neurais de regressão generalizada foram utilizadas. Os dados usados no treinamento e validação da rede foram obtidos através de um simulador rigoroso. Testes foram realizados e os autores concluíram que a retirada de algumas temperaturas de entrada não deteriora a predição do modelo, o que diminui o número de parâmetros necessários e melhora o tempo de resposta online. O sensor GRNN apresentou melhores resultados, mas ambos os sensores apresentaram performance viável para uso online, pois realizam boa predição e necessitam de baixo esforço computacional.

Com o objetivo de avaliar a qualidade do querosene para monitoramento e controle em tempo real de um processo de destilação de petróleo em uma CDU (*Crude Distillation Unit*), Sun *et al.* (2016) desenvolveram um sensor virtual baseado em redes neurais artificiais. Foram aplicadas técnicas de seleção de variáveis com o intuito de melhorar a precisão da predição do modelo, reduzir a sua complexidade e diminuir os custos das medições. Foi utilizada uma rede neural do tipo Perceptron de Multi-Camadas (MLP) e para avaliar as variáveis foi aplicado o método LASSO (*least absolut shrinkage and selection operator*) de análise de regressão. Para testar a sua eficiência, a rede LASSO-MLP foi comparada com as redes SBS-MLP e NMIFS, outras redes acopladas a técnicas de seleção de variáveis, e com uma rede MLP tradicional. O modelo proposto mostrou melhor desempenho que os outros modelos comparados.

Kataria *et al.* (2016) elaboraram um sensor virtual para inferir a composição dos produtos de fundo de uma coluna de destilação reativa. Para isto, foi estruturada uma rede neural artificial do tipo perceptron de múltiplas camadas com algoritmo Levenberg-Marquardt na otimização dos parâmetros. Os dados de treinamento e validação foram gerados a partir de

uma simulação rigorosa que considera que o processo segue o modelo UNIQUAC. O sensor apresentou resultados satisfatórios, com baixo MSE e alto R^2 .

Jalee e Aparna (2016) propuseram um sensor virtual para fazer a predição da composição dos produtos de topo e de fundo de um processo de destilação de benzeno e tolueno. O modelo proposto utilizou redes não lineares autorregressivas (NARX) baseadas em sistemas inferenciais neuro difusos adaptativos (ANFIS) na predição das frações molares dos produtos. Algoritmos de busca foram utilizados na determinação das variáveis de entrada e um simulador rigoroso na geração dos dados de treinamento. O sensor ANFIS apresentou bons resultados e mostrou-se uma alternativa viável aos analisadores de linha, geralmente utilizados na situação tratada pelo estudo.

Ito *et al.* (2018) modelaram um sensor virtual para determinar a composição de pentanos no topo de uma coluna debutanizadora utilizando um modelo de primeiro princípio. O modelo foi embasado nas equações que descrevem processos de destilação flash, combinam as condições de equilíbrio líquido vapor da mistura, balanços de massa e energia e equações de estado para determinar o coeficiente de fugacidade. Os autores utilizaram regras heurísticas fundamentadas no conhecimento do processo para determinar os componentes chave leve e pesado da mistura. Os resultados foram comparados com o de um sensor mais rigoroso e apresentou bons resultados apesar de considerar menos variáveis de entrada.

Jana e Banerjee (2018) construíram um sensor virtual que seria parte de um sistema de controle de uma coluna de destilação reativa. O objetivo do sensor virtual é inferir a composição do produto de fundo da coluna. Para isso, foi utilizado um modelo baseado em redes neurais que recebe as informações da temperatura online via termopar e estima a composição do produto de fundo para o controlador. A rede neural foi treinada utilizando o algoritmo Levenberg-Marquardt. Os autores chamaram o conjunto controlador-sensor virtual de controlador inferencial.

A partir do que vem sendo publicado nos últimos anos, nota-se que é vasto o número de trabalhos que aplicam sensores inferenciais em colunas de destilação. Na maior parte dos casos os autores utilizam métodos empíricos, como pode ser visto na Tabela 2, tendo como justificativa a complexidade do uso de modelos fenomenológicos, a sua alta demanda computacional e a facilidade em obter dados dos processos industriais. Porém, é comum que sejam utilizados dados gerados por simuladores robustos, que utilizam modelos fenomenológicos, na construção dos sensores virtuais de caixa preta.

Apesar do uso de Redes Neurais Artificiais na modelagem de colunas de destilação ser uma técnica consolidada, muitos trabalhos encontrados na literatura carecem de validação com dados reais. Utilizar dados gerados por simulações limita a predição do sensor à qualidade do modelo utilizado na geração do conjunto de dados. Um dos diferenciais deste trabalho consiste que o sensor foi calibrado e validado com dados reais, obtidos por experimentos em laboratório. Ademais, nota-se a escassez de publicações dessa natureza relacionadas à indústria de alimentos.

Colunas de destilação são exemplos de sistemas complexos de comportamento dinâmico e altamente não linear. De acordo com Abdullah *et al* (2007), a destilação apresenta um comportamento não linear devido as relações de equilíbrio líquido-vapor dos seus componentes e dinâmico por conta da rápida variação nas taxas de fluxo de vapor, lenta na temperatura e muito lenta na composição. Entre outras propriedades, as RNA possuem a capacidade de mapear funções não lineares, sendo muito utilizadas na modelagem de sensores virtuais. Elas podem aprender as relações não lineares entre as variáveis do sistema e fornecer uma boa aproximação a partir dos dados. Os fundamentos das Redes Neurais Artificiais serão apresentados na próxima seção.

Tabela 2 - Aplicações de sensores virtuais em processos de destilação.

Autores	Ano	Processo	Dados	Modelo
Oisovici e Cruz	2000	Destilação batelada	Simulados	Fenomenológico
Zamprogna <i>et al.</i>	2002	Destilação batelada	Simulados	Empírico - PLS
Fortuna <i>et al.</i>	2004	Coluna debutanizadora	Reais	Empírico - Redes Neurais
Zhang <i>et al.</i>	2004	Destilação azeotrópica	Reais	Empírico - Analise Regressiva
Osorio <i>et al.</i>	2008	Destilação batelada	Reais	Empírico - Redes Neurais
Smets <i>et al.</i>	2010	GPL e produção de estireno	Reais	Híbrido
Mohler <i>et al.</i>	2011	CDU	Reais	Empírico - Hammerstein - Wiener
Rani <i>et al.</i>	2011	Destilação reativa	Simulados	Empírico – Redes Neurais
Rogina <i>et al.</i>	2011	CDU	Reais	Empírico - Redes Neurais
Ramli <i>et al.</i>	2014	Coluna debutanizadora	Simulados	Empírico - Redes Neurais e PLS
Rahman <i>et al.</i>	2015	Destilação batelada	Simulados	Empírico – Redes Neurais e OLS
Togashov e Zmeu	2015	Destilação contínua	Simulados	Empírico - Equações Estruturais
Patil e Ghate	2015	Destilação multicomponentes	Simulados	Empírico - Redes Neurais
Sun <i>et al.</i>	2016	CDU	Simulados	Empírico - Redes Neurais
Kattaria <i>et al.</i>	2016	Destilação reativa	Simulados	Empírico - Redes Neurais
Jalee e Aparna	2016	Destilação de Benzeno	Simulados	Empírico - Logica Fuzzy
Ito <i>et al.</i>	2018	Coluna debutanizadora	Simulados	Fenomenológico
Jana e Banerjee	2018	Destilação reativa	Simulados	Empírico - Redes Neurais

3.3 REDES NEURAIIS ARTIFICIAIS

Segundo Silva *et al.* (2010), as Redes Neurais Artificiais (RNA) são modelos computacionais inspirados no sistema nervoso dos seres vivos capazes de adquirir e armazenar conhecimento. Também podem ser definidos como um conjunto de unidades de processamento, os neurônios artificiais, que são conectados entre si por matrizes de pesos sinápticos, as sinapses artificiais.

Davoudi e Vaferi (2018) definem as Redes Neurais como paradigmas inteligentes que, baseados no processo de aprendizagem dos seres humanos, são muito utilizados para modelar diversos fenômenos. Segundo Fukuda (1992), a Rede Neural Artificial tenta mimetizar o comportamento do cérebro por meio de modelos matemáticos. De acordo com Haykin (2001), as Redes Neurais Artificiais se assemelham ao comportamento do cérebro em dois aspectos:

- A capacidade de extrair conhecimento do ambiente onde estão inseridas através de um processo de aprendizagem;
- As conexões entre os neurônios, onde a Rede Neural, tal qual o cérebro, armazena o conhecimento adquirido.

Na literatura algumas propriedades importantes são atribuídas as Redes Neurais, tais como adaptabilidade, não linearidade, capacidade de aprendizagem, generalização, tolerância a falhas, facilidade de prototipagem, etc. (HAYKIN, 2001; SILVA *et al.*, 2010).

Devido essas características as Redes Neurais Artificiais possuem aplicações em diversas áreas, tais como reconhecimento de padrões, memória associativa, otimização, aproximação de funções, modelagem e controle de sistemas, processamento de imagens e classificação (MEIRELES e ALMEIDA, 2003; AHAMED e AKTHAR, 2016).

3.3.1 Neurônio biológico

O contato inicial com a teoria das Redes Neurais Artificiais passa pela compreensão do sistema nervoso e o funcionamento do neurônio biológico, a sua inspiração. Segundo Haykin (2001), o sistema nervoso pode ser visto como um sistema de três estágios. O cérebro é a rede neural que processa as informações e toma as decisões apropriadas, os receptores, que recebem os estímulos do ambiente e convertem em sinais elétricos que são enviados para o cérebro e os atuadores, que convertem os impulsos gerados pela rede neural em respostas.

O neurônio é a estrutura fundamental do sistema nervoso humano e funciona como um dispositivo de processamento de sinais. A célula neural consiste em um corpo celular ou soma, que possui os mesmos subcomponentes celulares que as outras células do corpo (tais quais membrana celular, citoplasma, núcleo, etc), uma única ramificação longa partindo do corpo celular chamada axônio e ramificações menores que saem do corpo celular e do axônio chamadas dendritos, como ilustrado pela Figura 3 (RUSSEL e NORVIG, 2003).

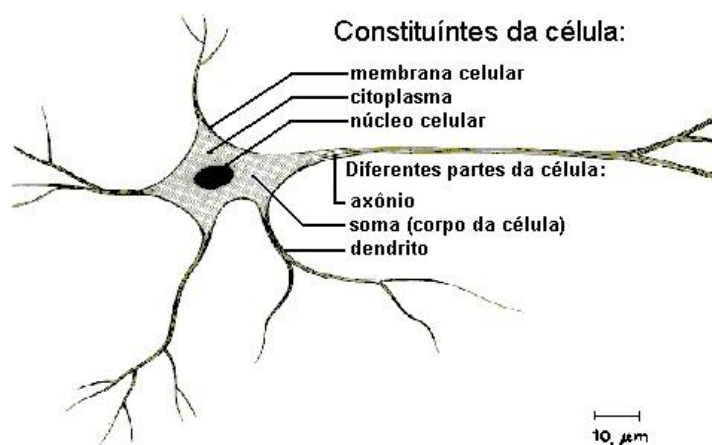


Figura 3 - Neurônio biológico. Fonte: ICMC-USP

Os neurônios estão conectados entre si com outros milhares de neurônios e se comunicam através de impulsos. Os impulsos vindos de outros neurônios atingem os dendritos e fluem até o corpo celular, onde são processados. Estes impulsos possuem a característica de “tudo ou nada”, ou seja, podem ser apenas excitatórios ou inibitórios. Caso o impulso recebido seja intenso o suficiente para excitar o neurônio, ou seja, maior que o limiar de ativação, ele desencadeia um potencial de ação que produz substâncias, chamadas neurotransmissores, que fluem através do axônio que pode ou não estar conectado a outros neurônios. O neurônio transmissor do impulso é capaz de controlar a sua frequência hiperpolarizando ou despolarizando a membrana pós-sináptica. São essas mudanças de polaridade que determinam a propagação do impulso nervoso. Enquanto o potencial de ação pode ser apenas excitatório ou inibitório, o potencial pós-sináptico possui uma gama de valores contínuos e depende da eficiência da sinapse em utilizar o neurotransmissor para produzir o impulso elétrico (GURNEY, 1997).

3.3.2 Neurônio Artificial

Como foi dito anteriormente, a Rede Neural Artificial tem como inspiração o sistema nervoso humano. Assim como o sistema nervoso humano tem como unidade fundamental o neurônio biológico, a unidade fundamental da Rede Neural Artificial é o neurônio artificial, mostrado pela Figura 4. De acordo com Krenker *et al.* (2011), o neurônio artificial consiste em um modelo matemático de três conjuntos de regras simples: multiplicação, soma e ativação. As entradas, (representadas aqui por $x_1, x_2, \dots, x_n$) são multiplicadas individualmente por pesos ($w_1, w_2, \dots, w_n$). No neurônio artificial, as entradas ponderadas são somadas com o bias (b) do

neurônio e processadas por uma função de ativação que determina o valor da sua saída. O modelo matemático do neurônio artificial pode ser expresso pelas Equações 1 e 2.

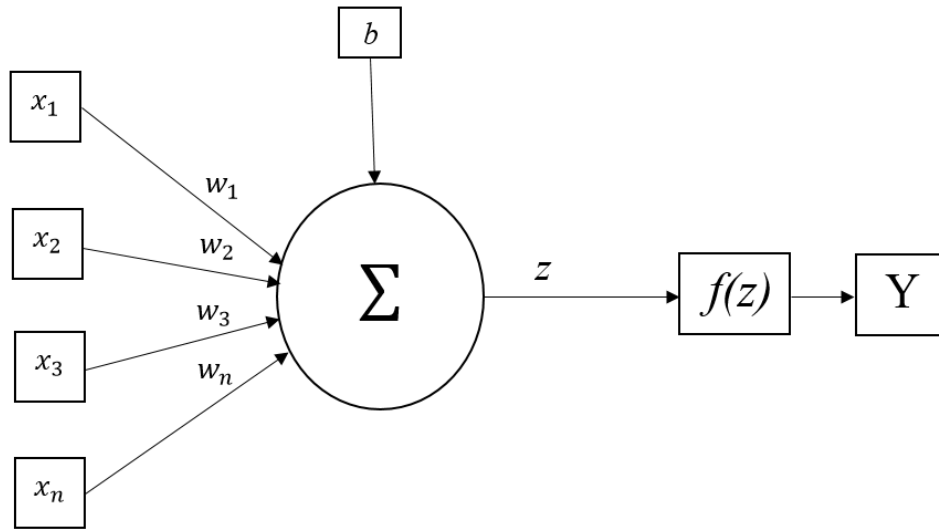


Figura 4 - Estrutura do neurônio artificial. Fonte: Elaborado pelo autor.

$$z = \sum_{n=1}^k w_n x_n + b \quad (1)$$

$$Y = f(z) \quad (2)$$

No qual Y é a saída fornecida pelo neurônio e f é a função de ativação. Segundo Hagan *et al.* (2014), as funções de ativação podem ser lineares ou não lineares em z e a sua escolha depende da natureza do problema que se deseja solucionar. Existem diversas funções que podem ser utilizadas como funções de ativação, algumas das mais utilizadas são citadas a seguir.

- Função limiar – Segundo exposto por Haykin (2001), a função limiar foi utilizada no modelo McCulloch-Pitts com o objetivo de simular a propriedade “tudo ou nada” do neurônio biológico. Neste modelo, a saída tem valor 1 caso o somatório das entradas ponderadas seja não negativo e valor 0 quando a soma é negativa. Hagan *et al.* (2014) ressaltam que esta função é bastante utilizada quando deseja-se construir uma Rede Neural que classifique as entradas em duas categorias distintas.

$$f(z) = \begin{cases} 1, & \text{se } z \geq 0 \\ 0, & \text{se } z < 0 \end{cases} \quad (3)$$

- Função linear – A função linear com $\beta = 1$ repete a saída do somador. Esta função é muito aplicada nas camadas de saída das Redes Neurais Artificiais do tipo Perceptron

de Múltiplas camadas e como função de ativação das redes do tipo ADALINE (KROSE e VAN DER SMAGT, 1996).

$$f(z) = \beta z \quad (4)$$

- Função logística – A função logística é uma função sigmoideal amplamente utilizada em Redes Neurais por ser contínua, monotônica e diferenciável em qualquer ponto (HAGAN *et al.*, 2014). A função logística possui natureza não linear e a sua saída fornece valores no intervalo $[0, 1]$ (FUKUDA, 1992).

$$f(z) = \frac{1}{1+e^{-\beta z}} \quad (5)$$

- Função tangente hiperbólica – Em relação à função logística, além de ser de natureza não linear, a função tangente hiperbólica possui a vantagem de preservar a forma sigmoideal e fornecer valores de saída positivos e negativos ($f(x) \in (-1, 1)$) (SILVA, 1998).

$$f(z) = \tanh(\beta z) \quad (6)$$

3.3.3 Arquitetura de uma Rede Neural

A maneira como os neurônios estão conectados entre si é o que define a arquitetura de uma Rede Neural. Desta forma, as principais arquiteturas são as Redes Neurais *feedforward* e as Redes Neurais recorrentes (KROSE e VAN DER SMAGT, 1996).

Neste trabalho foram utilizadas Redes Neurais do tipo *feedforward*. Também conhecidas como acíclicas ou de alimentação direta, nesta arquitetura a saída é calculada diretamente das entradas. Cada neurônio está conectado com os neurônios da camada seguinte e a informação é processada de forma direta, camada a camada, até gerar resposta na saída da rede (ANSARI *et al.*, 2018). As redes de alimentação direta podem apresentar estruturas com uma única camada, como as redes Perceptron e ADALINE, ou estruturas com múltiplas camadas, como as redes Perceptron de Múltiplas camadas e as Funções de Base Radial. As redes de única camada são utilizadas na resolução de problemas como reconhecimento de padrões e filtragem de problemas lineares, enquanto as redes de múltiplas camadas são mais utilizadas na resolução de problemas como aproximação de funções, classificação de padrões, identificação de sistemas, controle de processos, otimização, robótica, etc (SILVA *et al.*, 2016).

Diferente das redes *feedforward*, as redes recorrentes ou realimentadas apresentam comportamento dinâmico. Na arquitetura recorrente existe uma camada de contexto completamente conectada à camada escondida que utiliza o sinal prévio da camada de saída durante o treinamento (LOLLI *et al.*, 2017). Segundo Von Zuben (1996), a presença de realimentação possibilita a realização de representações internas e dispositivos de memória capazes de armazenar informações temporais e sinais sequenciais. Como resultado, a realimentação pode levar a comportamentos complexos mesmo com um número reduzido de parâmetros.

3.3.5 Aprendizado de uma Rede Neural

A principal propriedade de uma Rede Neural Artificial é a capacidade de aprender a partir de padrões que representam o comportamento do sistema e disponibilizar o conhecimento adquirido para a resolução de problemas. O objetivo do aprendizado é a generalização de soluções produzidas pela saída da rede, que devem representar de forma satisfatória o sistema que está sendo mapeado. O processo de aprendizagem consiste em ajustar os parâmetros e limiares da rede a partir de um conjunto de passos ordenados, denominado algoritmo de aprendizagem. Durante a aplicação de um algoritmo de aprendizado a Rede Neural absorve características específicas do sistema que se deseja mapear através dos seus dados (SILVA *et al.*, 2010). Para cada arquitetura de rede neural existem um ou mais algoritmos específicos e eles diferem entre si pela maneira com que os pesos são atualizados. Os paradigmas de aprendizagem podem ser divididos em dois, aprendizado supervisionado e aprendizado não supervisionado (KROSE e VAN DER SMAGT, 1996).

O aprendizado supervisionado é realizado por meio de exemplos compostos por um conjunto de dados rotulados, ou seja, a Rede Neural é exposta a um conjunto de dados das variáveis dependentes e independentes. Durante o treinamento, as variáveis independentes são propagadas pela rede que produz uma saída e o algoritmo de aprendizagem é aplicado para minimizar o erro entre a saída da rede e a saída desejada ajustando os valores dos parâmetros da rede neural (HASSANPOUR *et al.*, 2018). Conforme Russel e Norvig (2003), existem dois métodos de aprendizagem supervisionada, o aprendizado por classificação e o aprendizado por regressão. No aprendizado por classificação a saída fornece uma quantidade limitada de valores discretos e no aprendizado por regressão os valores produzidos pela saída são contínuos, reais.

No treinamento não supervisionado o conjunto de treinamento é formado por dados não rotulados, ou seja, a Rede Neural não recebe indicações sobre a saída desejada. Neste caso,

a meta é encontrar a estrutura subjacente dos dados originais (HUANG *et al.*, 2014). Neste paradigma, a RNA é treinada para fornecer uma saída que responda a padrões de agrupamento da entrada. Assim, deve-se descobrir características estatisticamente dominantes no conjunto de entrada. Ao contrário da aprendizagem supervisionada, como não há um conjunto de categorias a priori em que os padrões devem ser classificados, o sistema deve desenvolver sua própria resposta aos estímulos da entrada (KROSE e VAN DER SMAGT, 1996). Segundo Khosrowashahi (2011), os métodos de aprendizagem supervisionada utilizam iterações na minimização do erro ou da função objetivo enquanto os métodos não-supervisionados adotam uma abordagem analítica para determinação dos seus parâmetros.

3.3.6 Redes Perceptron de Múltiplas Camadas

Este trabalho utiliza uma rede neural do tipo Perceptron de Múltiplas Camadas para compor o sensor virtual. Consoante com Ansari *et al.* (2018) as Redes Perceptron de Múltiplas camadas são a arquitetura de Redes Neurais mais comumente utilizada. Segundo Hassanpour *et al.* (2018), as redes Perceptron de Múltiplas Camadas, também conhecidas como MLP, do inglês *multilayer perceptron*, consistem em uma rede neural de alimentação direta formada uma camada de entrada, uma ou mais camadas escondidas e uma camada de saída. As camadas escondidas podem ter números de neurônios distintos. O número de neurônios das camadas de entrada é obrigatoriamente o número de variáveis de entrada e saída que se deseja mapear. Nas redes MLP cada neurônio está conectado com os neurônios da camada posterior através dos pesos sinápticos e cada neurônio possui o seu próprio parâmetro ajustável, o bias.

As Redes diretas de múltiplas camadas são amplamente utilizadas na tarefa de aproximação de funções. De acordo com Cybenko (1989), uma rede MLP com uma camada escondida e função sigmoideal pode aproximar qualquer função arbitrária com precisão satisfatória desde que tenha o número de neurônios e o valor dos pesos adequado. Este teorema é conhecido como Teorema de Aproximação Universal.

3.3.7 Algoritmo de retropropagação do erro (*backpropagation*)

Seja uma Rede Neural Artificial arbitrária de n entradas e m saídas como a da Figura 5, onde:

- $X_1, X_2, \dots, X_n$ são os sinais de entrada;
- $W_{jn}^{(1)}, W_{lj}^{(2)}, W_{ml}^{(3)}$ são matrizes de pesos onde os elementos são valores dos pesos que conectam o j -ésimo neurônio da camada c ao n -ésimo neurônio da camada $c-1$;

- $b_j^{(1)}, b_l^{(2)}, b_m^{(3)}$ são matrizes de bias onde os elementos são os valores dos bias dos neurônios da camada c ;
- $z_j^{(c)}, z_l^{(c)}, z_m^{(c)}$ são vetores que representam as somas das entradas ponderadas do neurônio j da camada c ;
- $Y_j^{(c)}$ é a saída fornecida pelo j -ésimo neurônio da camada c ;

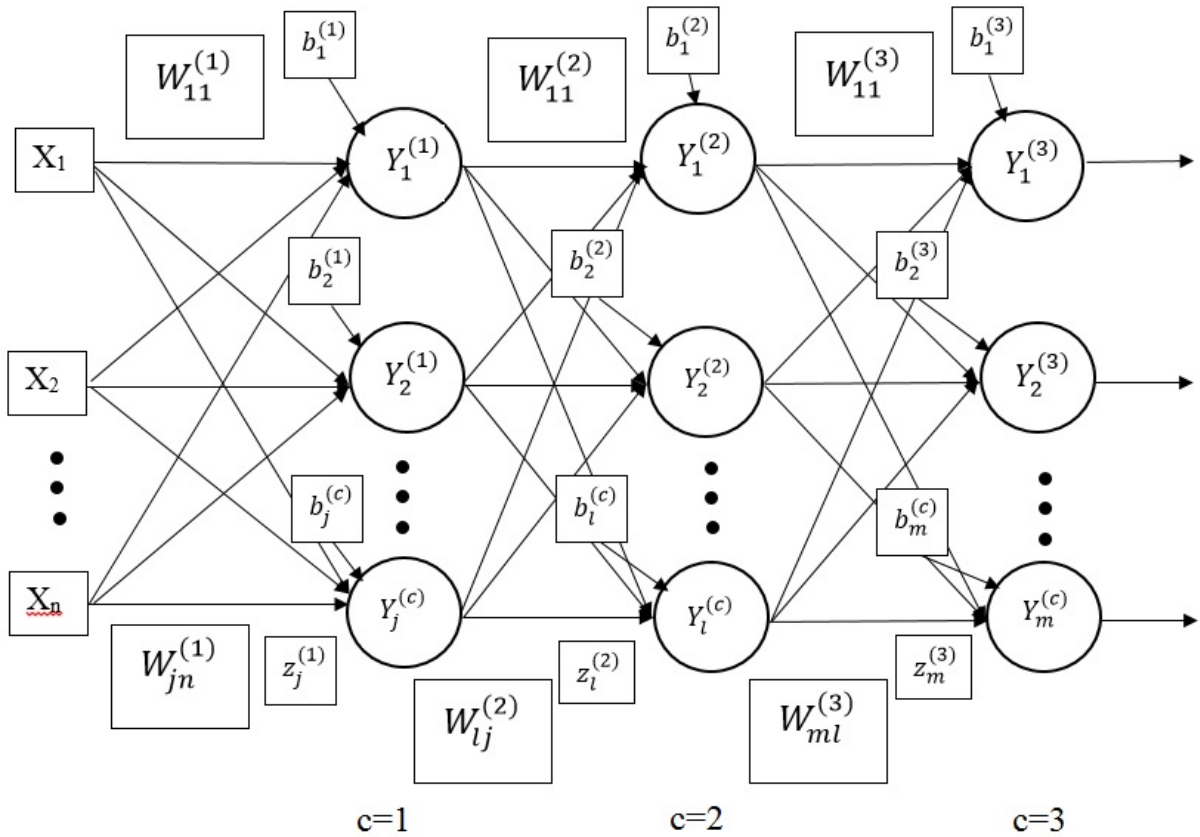


Figura 5 - Rede Neural Multicamadas. Fonte: Elaborada pelo autor.

Considerando a camada de saída da rede neural da Figura 5, tem-se que:

$$Y_m^{(3)} = f(z_m^{(3)}); \quad (7)$$

$$z_m^{(3)} = \sum W_{ml}^{(3)} Y_l^{(2)} + b_m^{(3)}; \quad (8)$$

Onde f é uma função contínua e diferenciável.

O algoritmo consiste em dois passos. No primeiro passo o sinal é propagado através da rede utilizando valores aleatórios dos pesos e bias e a rede fornece os respectivos valores de saída. Considerando a k -ésima amostra, a função erro que calcula a diferença entre a saída fornecida pela rede neural e a saída desejada é expressa pela Equação 9.

$$E(k) = \sum_{i=1}^m (\hat{Y}_i - Y_i^{(3)}) \quad (9)$$

No qual $\hat{Y}_i$ é o valor desejado e $Y_i^{(3)}$ o valor fornecido pela RNA. Para um conjunto de p amostras, o desempenho da rede neural é avaliado pelo Erro Quadrático Médio (EQM).

$$EQM = \frac{1}{p} \sum_{k=1}^p E(k) \quad (10)$$

A ideia do algoritmo *backpropagation* é encontrar valores dos pesos e bias da rede neural que minimizem a função erro, também chamada de função de custo, utilizando o método do gradiente descendente (ROJAS, 1996). Considerando a k -ésima amostra e o j -ésimo neurônio da camada de saída, aplicando a definição de gradiente na função erro e a regra da cadeia, tem-se que:

$$\nabla E = \frac{\partial E}{\partial W_{ml}^{(3)}} = \frac{\partial E}{\partial Y_m^{(3)}} \frac{\partial Y_m^{(3)}}{\partial z_m^{(3)}} \frac{\partial z_m^{(3)}}{\partial W_{ml}^{(3)}} \quad (11)$$

Onde:

$$\frac{\partial E}{\partial Y_m^{(3)}} = -(\hat{Y}_m - Y_m^{(3)}) \quad (12)$$

$$\frac{\partial Y_m^{(3)}}{\partial z_m^{(3)}} = f'(z_m^{(3)}) \quad (13)$$

$$\frac{\partial z_m^{(3)}}{\partial W_{ml}^{(3)}} = Y_l^{(2)} \quad (14)$$

Desta forma, tem-se:

$$\frac{\partial E}{\partial W_{ml}^{(3)}} = -(\hat{Y}_m - Y_m^{(3)}) f'(z_m^{(3)}) Y_l^{(2)} \quad (15)$$

O ajuste dos pesos é efetuado na direção contrária ao gradiente a fim de minimizar o valor da função erro, então:

$$\Delta W_{ml}^{(3)} = -\eta \frac{\partial E}{\partial W_{ml}^{(3)}} \quad (16)$$

$$\Delta W_{ml}^{(3)} = \eta \delta_m^{(3)} Y_l^{(2)} \quad (17)$$

$$\delta_m^{(3)} = f'(z_m^{(3)}) (\hat{Y}_m - Y_m^{(3)}) \quad (18)$$

No qual $\delta_m^{(3)}$ é o gradiente local em relação ao m -ésimo neurônio da camada de saída e η é a taxa de aprendizagem do algoritmo *backpropagation*. A regra de atualização dos pesos da camada de saída é expressa pela Equação 19.

$$W_{ml}^{(3)}(t+1) = W_{ml}^{(3)}(t) + \eta \delta_m^{(3)} Y_l^{(2)} \quad (19)$$

Onde o índice $(t+1)$ representa a matriz de pesos atualizada e (t) a matriz anterior. Assim, o algoritmo atualiza os parâmetros da camada de saída para a camada de entrada, camada a camada. A atualização dos bias é realizada de forma análoga.

Através da Equação 18, observa-se que o gradiente local $\delta_m^{(3)}$ depende dos valores de saída da Rede Neural, de forma que a Equação 18 não explica por si só a atualização das matrizes de pesos das camadas intermediárias. Neste sentido, segundo Haykin (2001), o gradiente local depende do neurônio j ser um neurônio de saída ou um neurônio oculto. Caso j seja um neurônio de saída, o gradiente é calculado pela Equação 18. Se j for um neurônio oculto, $\delta_j^{(c)}$ é igual ao produto da derivada pela soma ponderada dos δ s calculados para a camada posterior. Ou seja, os pesos das camadas de saída são ajustados primeiramente utilizando valores reais de erro, enquanto os pesos das camadas escondidas são atualizados utilizando os valores dos pesos já atualizados da camada de saída, que foram ajustados a partir do erro calculado na saída da rede neural, daí vem o nome “retropropagação”, visto que o erro é propagado da camada de saída até as camadas anteriores durante o ajuste.

3.3.8 Algoritmo Levenberg-Marquardt

De acordo com Silva *et al.* (2010), diversas variações do método *backpropagation* têm sido propostas a fim de melhorar a sua eficiência, sendo as mais difundidas a adição do termo *momentum*, o *resilient-backpropagation* e o algoritmo Levenberg-Marquardt. O presente trabalho utiliza o algoritmo Levenberg-Marquardt como algoritmo de aprendizado supervisionado para a rede MLP que compõe o sensor virtual que faz a predição da composição do etanol em tempo real.

O algoritmo Levenberg-Marquardt (LM) é baseado no método de Newton, técnica padrão para resolver problemas de mínimos quadrados não lineares. O método consiste em reduzir iterativamente a soma dos quadrados dos erros entre a função gerada e os dados através de uma sequência de atualizações nos valores dos parâmetros. Enquanto o algoritmo *backpropagation* é um método de gradiente descendente, o algoritmo LM é uma combinação

de dois métodos de minimização, o método do gradiente descendente e o método de Gauss-Newton e pode ser considerado uma extensão do método de Newton (HAGAN e MENHAJ, 1994; GAVIN, 2017).

A descrição dos passos do algoritmo que será dada a seguir, encontra-se de forma mais detalhada em Hagan e Menhaj (1994), Silva *et al.* (2010) e Hagan *et al.* (2014).

Assim como o algoritmo *backpropagation*, o algoritmo LM é executado em dois passos. No primeiro passo os sinais de entrada são propagados através da RNA e o erro entre a saída desejada e a saída fornecida pela Rede é calculado pela Equação 9. No mesmo sentido a avaliação do desempenho é fornecida pelo EQM, calculado pela Equação 10 e o objetivo aqui é encontrar valores de pesos e bias que minimizem a função erro. De acordo com Silva *et al.* (2010), as Equações 9 e 10 podem ser expressas de forma tal qual a Equação 20.

$$V = \frac{1}{2p} \sum_{k=1}^p E^T(k)E(k) \quad (20)$$

Como dito anteriormente, p é o número de amostras, k o seu respectivo índice e o sobrescrito T representa a transposta dos valores calculados para o erro. Seja w um vetor paramétrico representado pelos pesos, o Método de Newton minimiza $V(w)$ da forma descrita pela Equação 21.

$$\Delta w = -[\nabla^2 V(w)]^{-1} \nabla V(w) \quad (21)$$

Se $V(w)$ é uma soma de funções quadráticas, pode ser descrito pela Equação 22.

$$V(w) = \sum_{i=1}^N e_i^2(w) \quad (22)$$

E a partir da Equação 22, pode ser mostrado que:

$$\nabla V(w) = J^T(w)e(w) \quad (23)$$

$$\nabla^2 V(w) = J^T(w)J(w) + \mu I \quad (24)$$

O algoritmo LM atua como um método de segunda ordem, tal como um método quase-Newton, sem a necessidade de computar a matriz Hessiana, que é uma matriz de derivadas segundas (HAGAN *et al.*, 2014). Neste método, a matriz Hessiana é aproximada pela Equação 25.

$$H = J^T(w)J(w) \quad (25)$$

No qual H é a matriz Hessiana, J é a matriz Jacobiana, I é a matriz identidade e μ é um fator de ajuste do algoritmo. A matriz Jacobiana é expressa pela Equação 26.

$$J(w) = \begin{bmatrix} \frac{\partial e_1(w)}{\partial w_1} & \frac{\partial e_1(w)}{\partial w_2} & \frac{\partial e_1(w)}{\partial w_n} \\ \frac{\partial e_2(w)}{\partial w_1} & \frac{\partial e_2(w)}{\partial w_2} & \frac{\partial e_2(w)}{\partial w_n} \\ \frac{\partial e_n(w)}{\partial w_1} & \frac{\partial e_n(w)}{\partial w_2} & \frac{\partial e_n(w)}{\partial w_n} \end{bmatrix} \quad (26)$$

Observa-se que a matriz Jacobiana é da ordem do número de parâmetros da Rede Neural. Por conta disso, o seu cálculo para redes com grande número de conexões pode se tornar demasiadamente complicado. Hagan e Menhaj (1994), recomendam a utilização do algoritmo LM no treinamento de RNAs com até uma centena de parâmetros. A computação da matriz Jacobiana é a principal característica do algoritmo LM. Os elementos da matriz Jacobiana são calculados nas fases direta e inversa como no algoritmo *backpropagation* convencional (SILVA *et al.*, 2010).

Combinando as Equações 23 e 24 com a Equação 21, chega-se à regra de atualização do algoritmo Levenberg-Marquardt, descrito pela Equação 27.

$$\Delta w = -[J^T(w)J(w) + \mu I]^{-1}J^T(w)e(w) \quad (27)$$

Assim, o segundo passo do algoritmo diz respeito ao cálculo do ajuste do valor dos pesos que minimiza $V(w)$. Durante o treinamento, o algoritmo controla a direção no espaço de erros através do fator de ajuste μ . Cada vez que os pesos são atualizados, analisa-se se a atualização aumenta ou diminui o valor de $V(w)$. Sempre que a atualização aumenta o valor de $V(w)$, μ é multiplicado por um parâmetro β . Quando o passo de atualização diminui o valor de $V(w)$, μ é dividido por β . Desta maneira, para fatores de ajuste altos, o algoritmo LM aproxima-se do método do gradiente descendente e para fatores de ajuste baixos do método de Gauss-Newton (HAGAN e MENHAJ, 1994).

4 METODOLOGIA

O sensor virtual proposto tem como variável de entrada a temperatura do prato de topo e como saída a fração mássica do etanol. No seu desenvolvimento, o trabalho foi dividido em duas etapas, experimental e computacional. Na etapa experimental foram realizadas as destilações de misturas binárias etanol-água que geraram o conjunto de dados utilizado no desenvolvimento do sensor virtual, além da produção do mosto fermentado e as destilações utilizadas em testes posteriores. A etapa computacional consiste no desenvolvimento da Rede Neural Artificial que compõe o sensor virtual.

4.1 ETAPA EXPERIMENTAL

4.1.1 Materiais utilizados

As destilações foram realizadas no Laboratório de Engenharia de Sistemas em Alimentos e Inovação do Departamento de Tecnologia de Alimentos da Universidade Federal de Sergipe. O equipamento utilizado consiste em um módulo experimental de destilação da Eco Educacional (Figura 6) que opera em batelada. A coluna é totalmente transparente e possui isolamento térmico a vácuo. Possui sistema de coleta de amostras com sete módulos encamisados de vidro/Teflon com diâmetro interno de 65 mm e altura de 100 mm, preenchidos com anéis de Rasching de borosilicato com 8 mm de diâmetro. Estão distribuídos pelo equipamento dez sensores PT-100 para fins de monitoramento de temperatura, sendo um na saída do destilado, um no prato de topo, sete ao longo da coluna (um em cada módulo de coleta) e um no refeedor, cuja leitura é realizada por meio de painel eletrônico. A coluna possui altura de 90 cm e acima dela está o condensador conectado a um banho ultratermostático SL 152 com temperatura de 10°C para reduzir as perdas durante as destilações. Abaixo da coluna está o refeedor, que é aquecido através de uma manta elétrica e possui capacidade de 5 L. A potência do equipamento é de 1000 W.

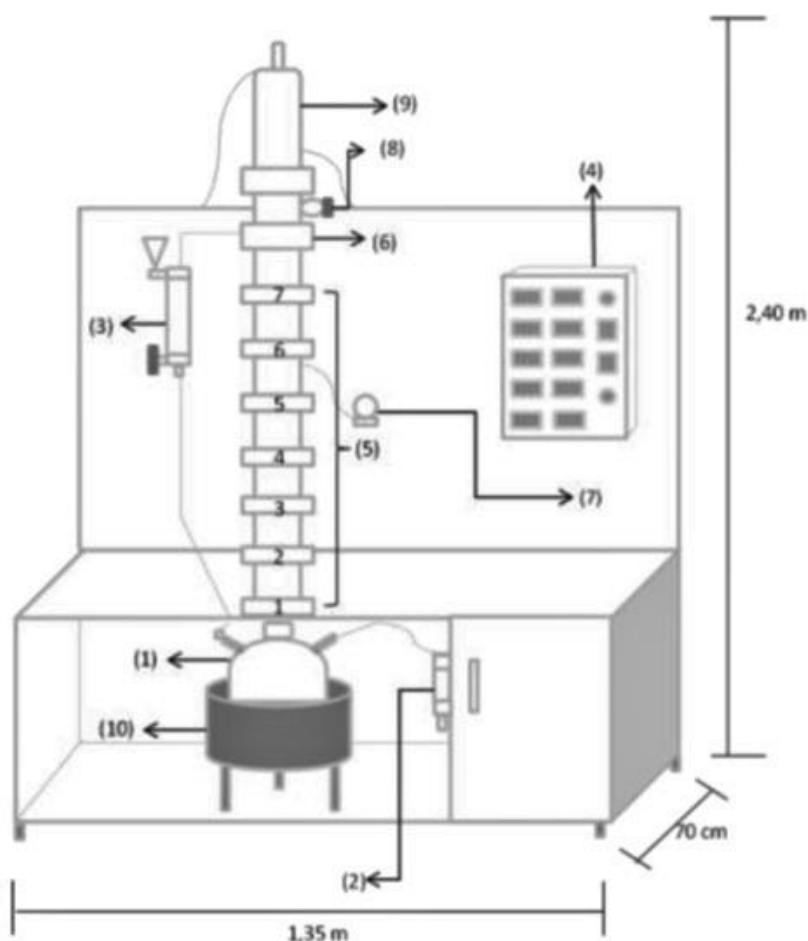


Figura 6 - Módulo experimental de destilação. (1) – Refervedor, (2) – Coletor de amostra (presente no refervedor), (3) – Coletor de amostra (produto), (4) – Painel eletrônico, (5) – Módulos de coletas (pratos), (6) – Módulo de topo, (7) – Vacuômetro Bourdon, (8) – Refluxador, (9) – Condensador e (10) – Manta de aquecimento. Fonte: Mota (2017).

4.1.2 Destilações das misturas sintéticas

Com o objetivo de gerar um conjunto de dados suficiente para um treinamento eficiente da Rede Neural Artificial que compõe o sensor, seis destilações de misturas binárias etanol-água com concentrações de etanol iniciais próximas das concentrações dos mostos fermentados produzidos em laboratório foram realizadas.

As misturas consistem em água destilada e álcool etílico, com concentrações iniciais de 10,4%, 12,4%, 12,7%, 12,7%, 13,1% e 13,4%(v/v) e foram produzidas utilizando álcool etílico hidratado (96°GL) e água destilada. Quantidades específicas de água e álcool foram misturadas a fim de produzir misturas binárias de concentrações iniciais próximas das concentrações dos mostos fermentados. Tais concentrações foram determinadas manualmente por meio de um densímetro digital do fabricante *Kyoto Electronics Manufacturing*, modelo

DA-130 N, com precisão de $\pm 0,001 \text{ g/cm}^3$. No início do experimento, o aquecimento do refrator foi ajustado em 700 W e após a sua temperatura ter atingido 50°C, reajustado em 750 W. Durante todo o processo a coluna operou em isolamento térmico a vácuo de 600 a 650 mmHg e razão de refluxo 1:2. A razão de refluxo 1:2 consiste em enviar o vapor condensado para ser coletado durante 10s e devolvê-lo para a coluna durante 5s. Todos os experimentos tiveram volume inicial de 4,5 L.

Durante os experimentos as amostras foram coletadas manualmente a cada 2 minutos e tiveram o seu teor alcóolico mensurado também por meio de densímetro digital. À medida que as amostras foram coletadas, as temperaturas foram monitoradas por meio de painel eletrônico. O processo seguiu até que o teor de álcool nas amostras foi considerado insignificante.

4.1.3 Produção do mosto fermentado

A capacidade de generalização em situações reais foi verificada com a produção do mosto fermentado de melancia. A sua produção seguiu a metodologia utilizada por Costa (2016). Para a produção do mosto foram adquiridas em mercado local 4 unidades de melancia (*citrullus lanatus*) que totalizaram 28,815 kg, dos quais 15,668 kg de polpa, 10,820 kg de cascas e 618 g de sementes. Após a etapa de filtração, foram gerados 13,7 L de suco e 1,556 g de resíduos ficaram retidos na peneira. Ao suco foram adicionados 6,9 L de água e 3,139 kg de açúcar para ajuste do teor de sólidos solúveis, que iniciou o processo de fermentação em 17,9 °Brix. A levedura utilizada foi a *saccharomyces cerevisiae*, adquirida em mercado local. A fermentação ocorreu em garrafas de 5 L adaptadas com mangueiras para a liberação de dióxido de carbono e durou seis dias, sendo encerrada quando o teor de sólidos solúveis (°Brix) apresentou valores constantes. Ao final do processo foram produzidos 22,350 L de mosto fermentado de melancia, que após a etapa de trasfega para retirada de sedimentos e recuperação da levedura foram utilizados na destilação para a produção da bebida destilada de melancia. A concentração inicial de etanol do mosto obtido foi de 11%(v/v).

4.1.4 Destilação do mosto fermentado e produção da aguardente de fruta

De posse do mosto fermentado de melancia foram executadas as destilações que geraram a bebida destilada. Foram realizadas quatro bateladas, três com volume inicial de 4,5 L e a última de 4 L. O processo seguiu as mesmas condições operacionais das destilações sintéticas e as amostras foram coletadas a cada 5 minutos até que o seu teor de álcool fosse

desprezível. Dados de três das quatro destilações foram utilizados devido problemas operacionais que ocorreram durante uma das bateladas, quando foi necessário interromper o experimento antes do programado. Na Tabela 3 encontram-se as informações relativas ao rendimento das destilações do mosto fermentado de melancia utilizadas na validação do sensor virtual.

Tabela 3 - Rendimento das destilações do mosto fermentado de melancia.

	Destilação 1		Destilação 2		Destilação 3	
	Volume	Teor alcoólico	Volume	Teor alcoólico	Volume	Teor alcoólico
Cabeça	260 mL	87,2 %(v/v)	380 mL	86,2 %(v/v)	365 mL	87,2 %(v/v)
Coração	560 mL	44,4% (v/v)	370 mL	42,6% (v/v)	435 mL	41,9%(v/v)
Cauda	245 mL	4,8 %(v/v)	200 mL	6,4 %(v/v)	185 mL	5,2%(v/v)

Além das destilações do mosto fermentado de melancia foram utilizados dados de destilações realizadas em estudos anteriores do Laboratório de Engenharia de Sistemas em Alimentos e Inovação a fim de testar a capacidade de inferência do sensor. Os dados são relativos às quatro bateladas de um mosto fermentado de jabuticaba, de teor alcoólico inicial de 13,3% (v/v), utilizados no estudo de Mota (2017) e três bateladas de um mosto fermentado de manga, de Oliveira *et al.* (2018), de teor alcoólico inicial de 10,0 %(v/v), todas com volume inicial de 4,5 L e realizadas nas mesmas condições operacionais que a destilação do fermentado de melancia.

4.2 ETAPA COMPUTACIONAL

O sensor virtual proposto consiste em uma Rede Neural Artificial de alimentação direta, cuja arquitetura escolhida foi o Perceptron de Múltiplas Camadas com uma camada escondida. Todas as simulações foram implementadas em ambiente MATLAB, versão R2016a. Através das destilações das misturas binárias foi formado o conjunto de dados utilizado na construção da Rede Neural, que consiste em 350 amostras dos pares ordenados temperatura e teor de etanol. Tais conjuntos foram divididos em três subconjuntos de forma aleatória, sendo 70% dos dados utilizados para treinamento (conjunto de 245 amostras) e 15% para validação (52 amostras) e 15% para teste (52 amostras).

Com o objetivo de evitar o sobre treinamento ou *overfitting*, situação em que a rede está superajustada aos dados de treinamento e apresenta uma capacidade de generalização pobre, um dos critérios de parada do treinamento da Rede Neural foi a técnica de validação cruzada. A técnica de validação cruzada consiste em antecipar o término do treinamento através

da avaliação do erro no conjunto de validação. A rede é treinada apenas com o conjunto de treinamento e avalia-se o erro do conjunto de validação, que ainda não foi apresentado à rede, a cada k iterações (épocas). Como visto na seção 3.3.8, a cada iteração o erro é calculado e o valor dos pesos, ajustado. Se o erro calculado para o conjunto de dados de validação for maior que o calculado há k iterações, o treinamento é dado como encerrado e os valores dos pesos encontrados k iterações atrás são utilizados como resposta (SILVA, 1998).

Neste trabalho utilizou-se k igual a 6 épocas para o procedimento de validação cruzada. O treinamento também foi dado como encerrado caso o conjunto de validação atingisse o erro MSE igual ou menor que 10^{-4} , o valor da matriz Hessiana ultrapassasse 10^6 e o treinamento ocorreu por um número máximo de mil épocas.

O algoritmo de aprendizagem utilizado foi o algoritmo Levenberg-Marquardt, descrito na seção 3.3.8. Os pesos foram iniciados com valores aleatórios no intervalo $[-1,1]$ e o fator de ajuste foi fixado em 10. Foram testadas duas funções de ativação não lineares na camada escondida, as funções tangente hiperbólica e logística. Em todas as arquiteturas uma função de ativação linear foi utilizada na camada de saída. A Figura 7 ilustra o processo de treinamento da Rede Neural Artificial.

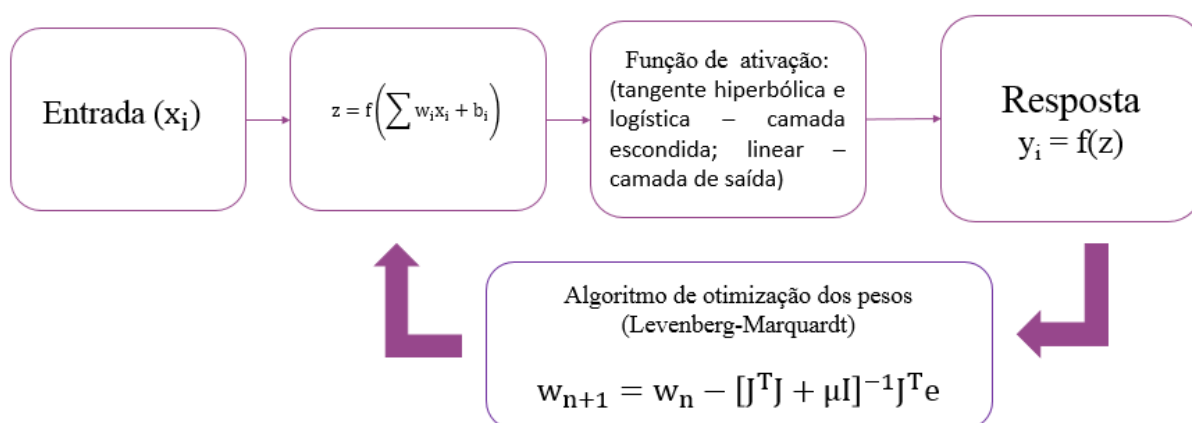


Figura 7 - Esquema de treinamento da RNA. Fonte: Elaborada pelo autor.

A determinação do número ótimo de neurônios foi feita por tentativa e erro via algoritmo construtivo, no qual partindo de um número mínimo são adicionados neurônios na camada escondida até que não seja observado ganho de desempenho com a adição de mais neurônios. Desta maneira, cada arquitetura foi testada 10 vezes variando de um a quinze neurônios na camada escondida, quando foi observado não haver ganho de performance com a adição de mais neurônios.

A Rede Neural escolhida foi a que apresentou o menor erro quadrático médio (EQM) quando apresentada a todos os dados, ou seja, a rede que melhor generaliza o conjunto de dados das destilações de água e álcool. A Equação 28 calcula o valor do EQM.

$$EQM = \frac{1}{N} \sum_{i=1}^N (y_{observado} - y_{estimado})^2 \quad (28)$$

Sendo N o número de amostras observadas.

A finalidade da construção do sensor é inferir a composição de etanol em tempo real em processos de destilações de mostos fermentados. Com este objetivo, o sensor desenvolvido teve a sua capacidade de inferência testada em relação às destilações de mostos de melancia, manga e jabuticaba. Para avaliar a sua performance, foram utilizados o erro absoluto médio (EAM) e o desvio padrão do erro (DP). As Equações 29 e 30 descrevem o erro absoluto médio e o desvio padrão do erro, respectivamente.

$$EAM = \frac{1}{N} \sum_{i=1}^N |y_{observado} - y_{estimado}| \quad (29)$$

$$DP = \sqrt{\frac{\sum_{i=1}^N (e_i - EAM)^2}{N}} \quad (30)$$

Sendo e_i é o erro de predição para a amostra i .

5 RESULTADOS E DISCUSSÕES

5.1 FORMAÇÃO DO CONJUNTO DE DADOS

Primeiramente foram analisadas as variáveis do processo que poderiam influenciar na variável que se deseja inferir, que no caso do presente problema é a composição do etanol. O problema possui como variáveis independentes a taxa de refluxo, temperatura do condensador, temperaturas em diversos estágios da coluna e a pressão na coluna. Exceto as temperaturas ao longo da coluna, todas as outras variáveis independentes se mantem constantes durante o processo, restando apenas as temperaturas como variáveis possivelmente correlacionadas com a composição do etanol no destilado. As Figuras 8-13 a-b ilustram o perfil de temperatura da coluna e da concentração do etanol no destilado durante as seis destilações da mistura binária. Os módulos estão indicados da seguinte maneira: os índices vão de 1 a 7, sendo 1 na base da coluna, mais próximo do refeedor e 7 mais próximo do estágio de topo.

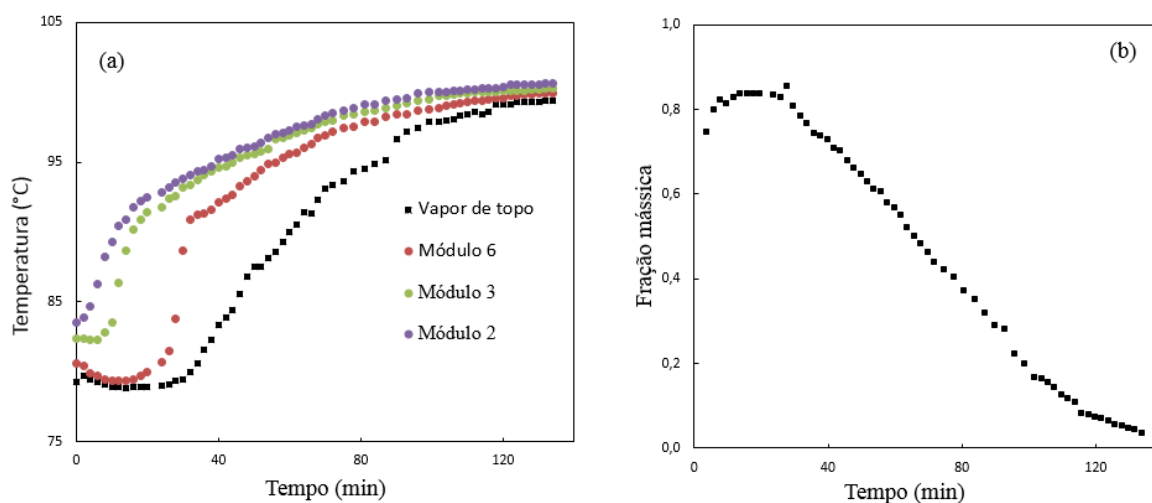


Figura 8 - Perfil de temperatura da coluna (a) e de etanol de concentração de etanol no destilado (b) durante o processo - Experimento 1. Concentração inicial: 12,7%(v/v).

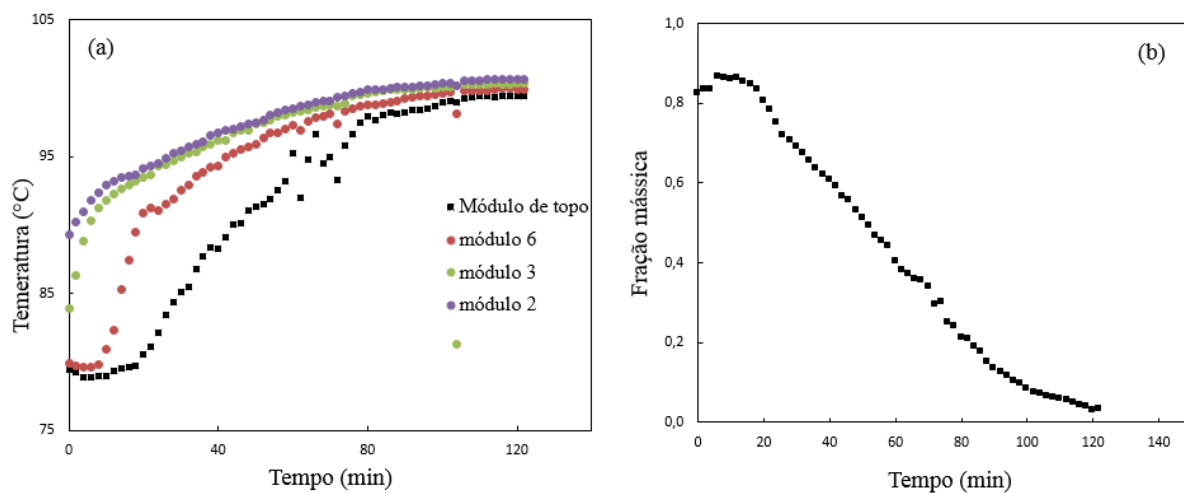


Figura 9 - Perfil de temperatura da coluna (a) e de concentração de etanol no destilado (b) durante o processo - Experimento 2. Concentração inicial: 13,1 %(v/v).

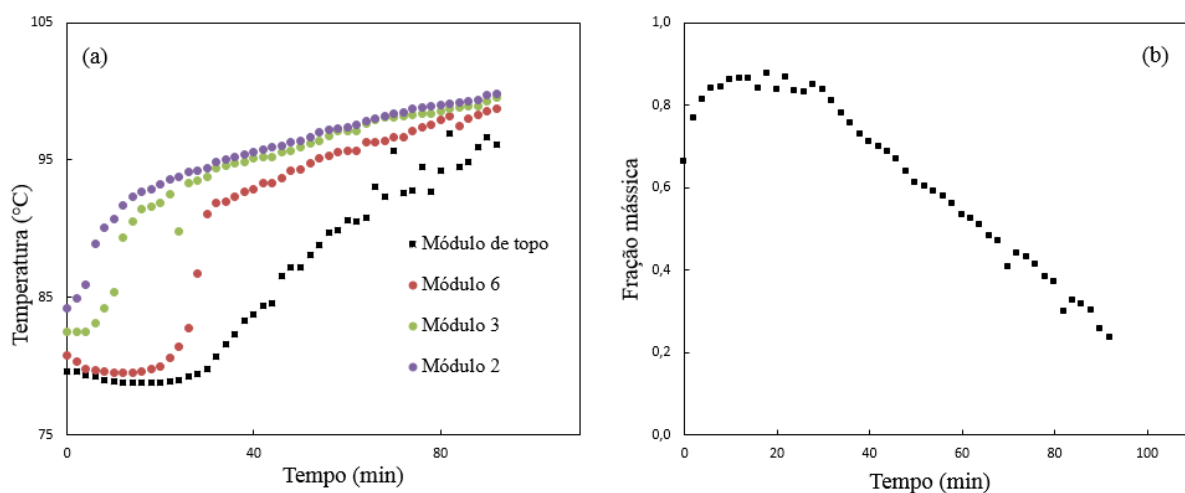


Figura 10 - Perfil de temperatura da coluna (a) e de concentração do etanol no destilado (b) durante o processo - Experimento 3. Concentração inicial: 10,4%(v/v).

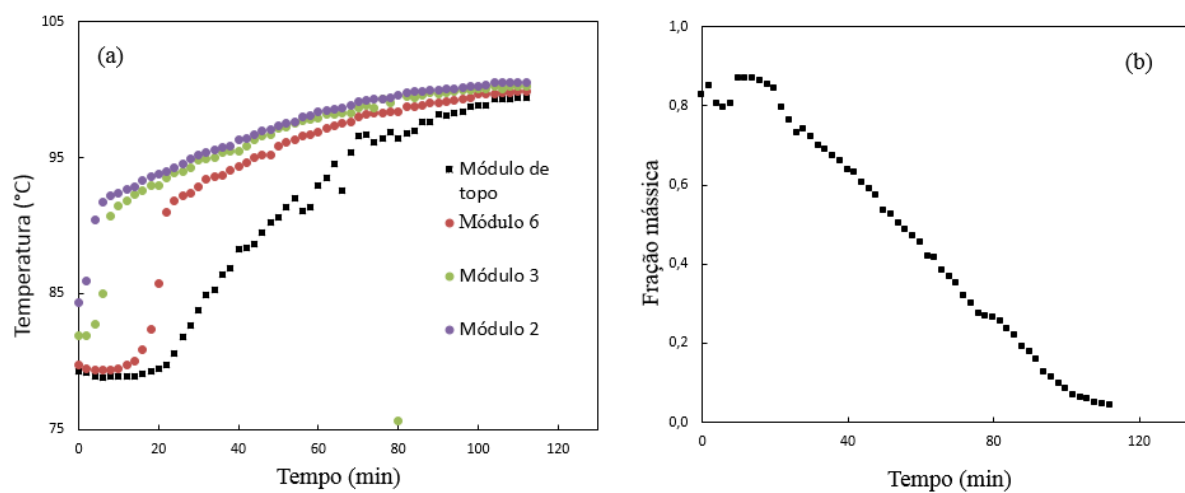


Figura 11 - Perfil de Temperatura da coluna (a) e de concentração do etanol no destilado (b) ao longo do processo - Experimento 4. Concentração inicial: 12,7 %(v/v).

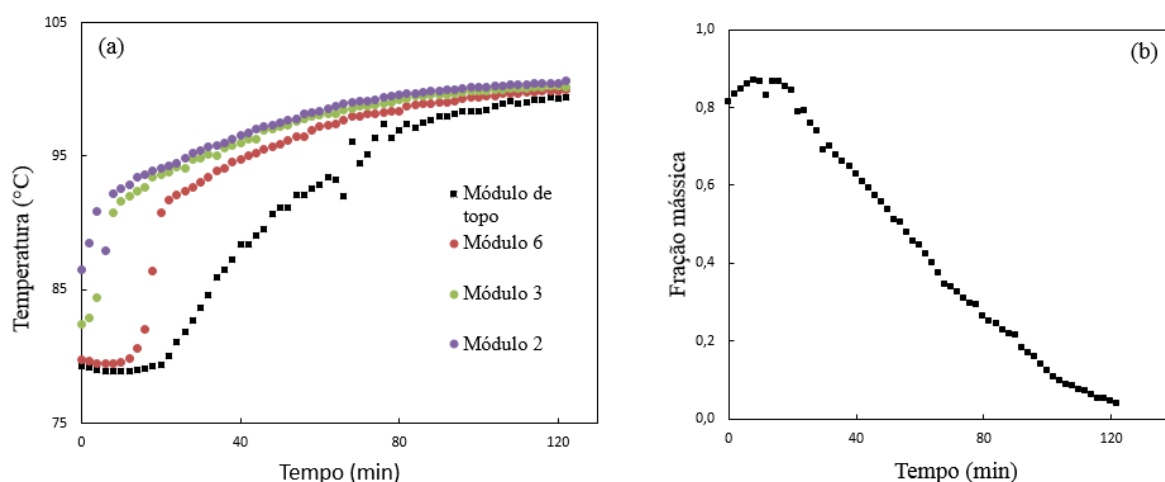


Figura 12 - Perfil de temperatura da coluna (a) e de concentração de etanol no destilado (b) durante o processo- Experimento 5. Concentração inicial: 13,4 %(v/v).

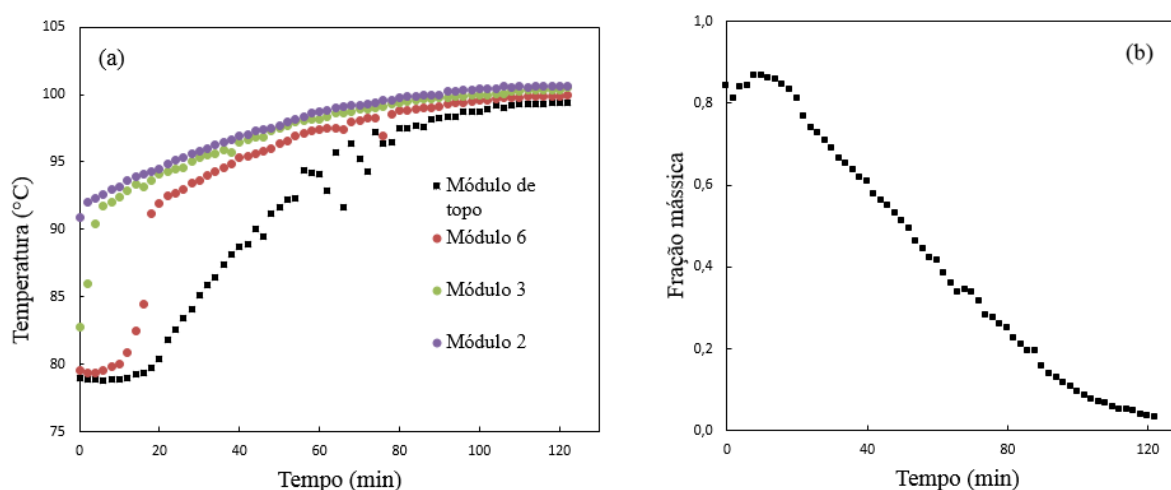


Figura 13 - Perfil de temperatura da coluna (a) e de concentração de etanol no destilado durante o processo - Experimento 6. Concentração inicial: 12,4 %(v/v).

Através das Figuras 8-13.b observa-se o perfil dinâmico da concentração durante o processo, esperado para destilações em batelada. O mesmo pode ser observado para as temperaturas através das Figuras 8-13.a. Também podem ser observadas perturbações na temperatura do estágio de topo, não identificadas nos perfis de temperatura dos módulos 2, 3 e 6. Isto acontece por conta de variações na temperatura do condensador. Em alguns momentos durante o processo foi necessário repor a água do banho termostático conectado ao condensador para que a sua temperatura fosse mantida em 10°C. Devido à proximidade do módulo de topo com o condensador, perturbações na temperatura do condensador acarretam em oscilações na temperatura do estágio de topo. Nota-se também que as temperaturas dos módulos 2, 3 e 6 atingem valores próximos da temperatura final do processo muito mais cedo que a temperatura do módulo de topo. Por conta dessas semelhanças entre o perfil dinâmico da temperatura do

vapor de topo e do perfil de concentração do etanol, apenas a temperatura de topo foi escolhida como variável de entrada. Além disso, em uma modelagem da mesma coluna de destilação, Mota (2017) utilizou apenas temperatura do módulo de topo com sucesso.

Identificadas as variáveis, iniciou-se o processo de construção da Rede Neural que deverá compor o sensor virtual. A Rede proposta tem uma entrada, que é o valor da temperatura no módulo de topo, e uma saída, a respectiva fração mássica de etanol no destilado. Assim, o conjunto de dados utilizados na representação do sistema consistem nos pares ordenados temperatura do módulo de topo e fração mássica de etanol e possui 350 amostras relativas aos dados experimentais.

5.2 CONSTRUÇÃO DA REDE NEURAL ARTIFICIAL

Esta etapa trata da determinação da configuração de Rede Neural que melhor generaliza os dados das destilações de água e álcool. Após treinadas, as Redes Neurais tiveram o seu desempenho analisado sendo apresentadas a todos os dados. A Tabela 4 mostra as 10 Redes Neurais de melhor desempenho, bem como o seu número de neurônios na camada escondida e a respectiva função de ativação. Os índices se referem ao Erro Quadrático Médio relativo aos conjuntos de treinamento (MSE_{tr}), validação (MSE_{val}), teste (MSE_{teste}) e ao desempenho da RNA quando apresentada a todos os dados (MSE_{tot}).

Tabela 4 - Desempenho das melhores Redes Neurais.

Índice	Número de neurônios	Função de ativação	MSE_{tr} ($\times 10^{-4}$)	MSE_{val} ($\times 10^{-4}$)	MSE_{teste} ($\times 10^{-4}$)	MSE_{tot} ($\times 10^{-4}$)
1	12	Logística	6,32	11,00	5,10	6,83
2	4	Tangente hiperbólica	5,88	8,62	9,48	6,84
3	8	Tangente hiperbólica	7,18	5,34	6,72	6,83
4	9	Tangente hiperbólica	6,77	3,19	11,00	6,79
5	11	Tangente hiperbólica	7,91	3,54	5,24	6,85
6	12	Tangente hiperbólica	6,85	7,54	5,68	6,81
7	13	Tangente hiperbólica	6,48	2,96	12,00	6,78
8	14	Tangente hiperbólica	6,40	11,00	5,05	6,85
9	14	Tangente hiperbólica	5,79	11,00	7,77	6,84
10	15	Tangente hiperbólica	6,45	11,00	4,37	6,83

Analisando as informações da Tabela 4, observa-se que as configurações com a função de ativação tangente hiperbólica na camada escondida apresentaram melhor desempenho que as configurações com função de ativação logística. As RNA mostradas pela Tabela 4 apresentaram valores muito próximos e todas podem ser utilizadas como sensor virtual. Devido a diferença de velocidade de processamento para RNA dessa magnitude ainda não ser um fator

determinante, o Erro Quadrático Médio foi utilizado como critério para a escolha da RNA em detrimento de outros fatores, como o número de neurônios na camada escondida. Neste sentido, a RNA de índice 7 foi escolhida para compor o sensor virtual por ter obtido o menor MSE quando apresentada a todos os dados. As Figuras 14 e 15 mostram o coeficiente de regressão R da RNA 7, bem como a sua curva de treinamento.

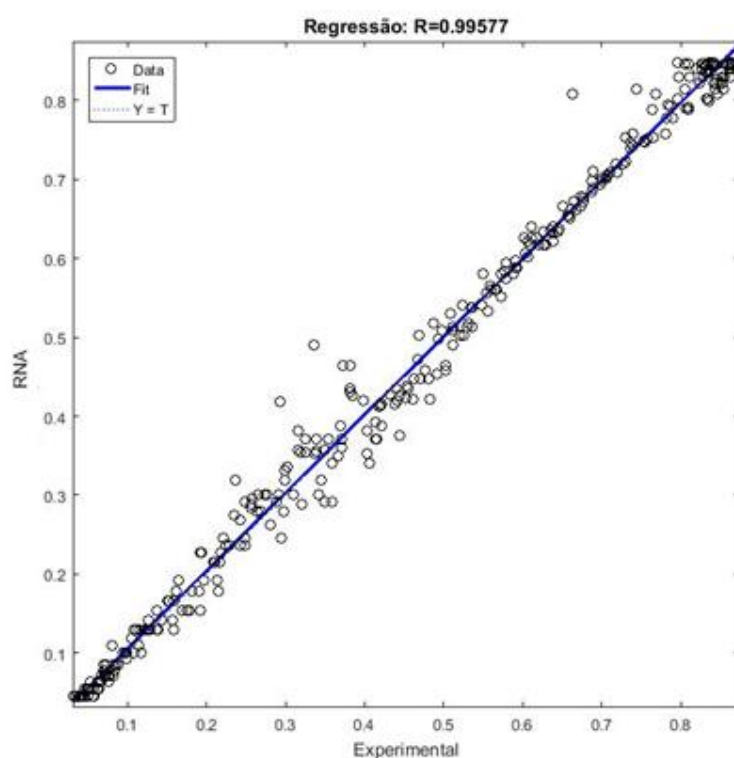


Figura 14 - Regressão da melhor topologia obtida.

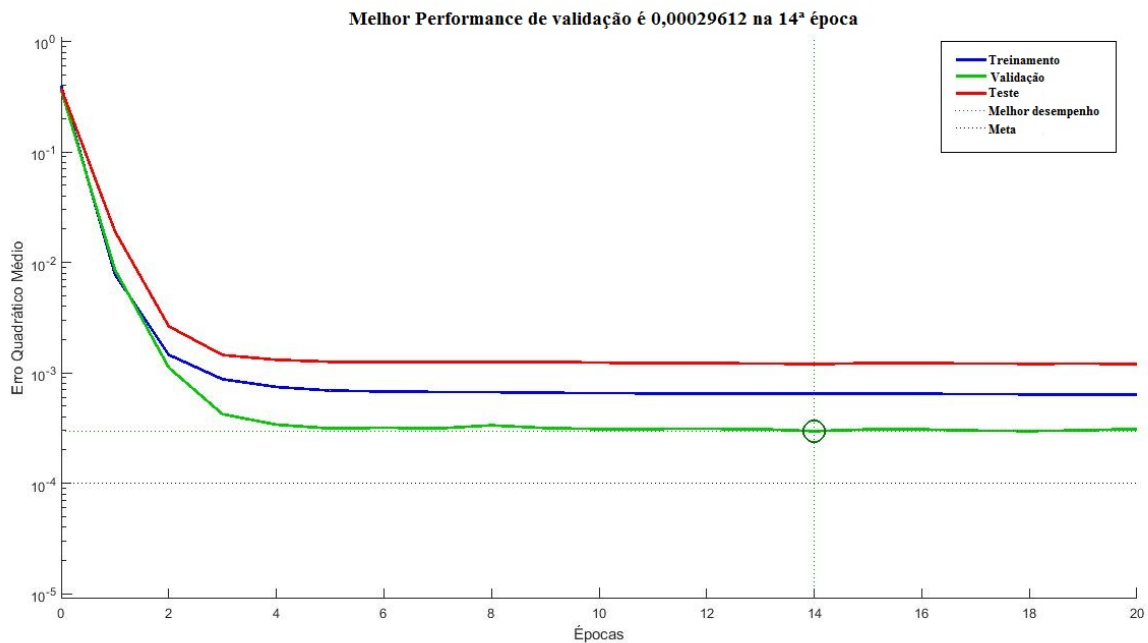


Figura 15 - Curva de treinamento da Rede Neural de melhor desempenho.

Como pode ser observado na Figura 13, a RNA desenvolvida realizou uma ótima representação dos dados experimentais, com R muito próximo de 1. Na Figura 14, nota-se que as curvas de treinamento, validação e teste seguem a mesma tendência após o critério de parada do treinamento, mostrando a não ocorrência de *overfitting*. Caso os dados estivessem excessivamente ajustados ao conjunto de treinamento, as curvas tomariam direções distintas e apresentariam um valor de erro crescente para o conjunto de validação. Como dito anteriormente, a validação cruzada foi utilizada como critério de parada do treinamento. O treinamento foi realizado em 20 iterações ou épocas e o erro entre os conjuntos de treinamento e validação começou a crescer a partir da 14ª iteração, sendo nesta época definidos os valores dos pesos e bias da RNA. A Figura 16 e a Tabela 5 expõem a topologia da Rede Neural e os valores dos seus parâmetros.

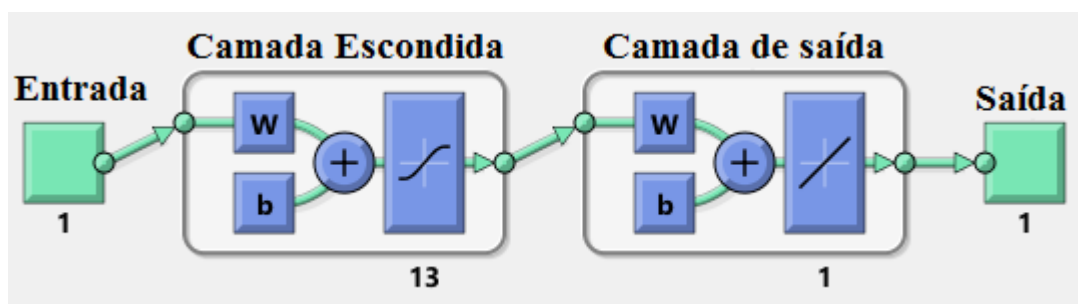


Figura 16 - Topologia da Rede Neural de melhor desempenho.

Tabela 5 - Valores dos pesos e bias da Rede Neural.

Neurônio	Camada escondida		Camada de saída	
	$W_{1,j}$	b_j	$W_{i,1}^T$	b
1	-18,578	20,61226	1,70673	-1,70602
2	10,76062	-9,30526	-0,28786	
3	20,94482	-14,1066	-0,06402	
4	-19,6413	9,825068	0,070943	
5	10,28892	-2,68178	-0,15543	
6	-19,4828	1,662724	0,049474	
7	-22,7459	-1,16169	0,053633	
8	-21,3581	-4,58053	0,044964	
9	18,30608	6,85372	-0,03759	
10	17,80444	9,843626	-0,03995	
11	18,80867	11,68038	-0,01712	
12	-19,3157	-14,7042	0,051011	
13	-28,096	-26,4213	0,0679	

Como mostrado pela Figura 16, a Rede Neural que melhor representa o comportamento das destilações da mistura binária possui um neurônio na camada de entrada, que recebe o valor da temperatura, uma camada escondida com 13 neurônios e função de ativação tangente hiperbólica e um neurônio na camada de saída com função de ativação linear, que fornece o valor da fração de etanol correspondente.

Através da Tabela 5 nota-se que a RNA possui um total de 40 parâmetros relativos às suas conexões, sendo 13 pesos que conectam a camada de entrada à camada escondida, 13 valores de bias, um para cada neurônio da camada escondida, 13 pesos que conectam a camada escondida à camada de saída e um valor de bias relativo ao neurônio da camada de saída. Utilizando os valores dos parâmetros da Tabela 5 é possível construir a Rede Neural que calcula o valor da fração de etanol no destilado durante o processo a partir de medidas da temperatura do estágio de topo.

Através das Figuras 17 a Figura 22 pode-se analisar a capacidade de inferência do sensor virtual desenvolvido durante os processos de destilação das misturas sintéticas. Na Tabela 6 constam os valores do erro absoluto médio, desvio padrão do erro, erro máximo e erro mínimo, utilizados na análise da qualidade da inferência fornecida pelo sensor.

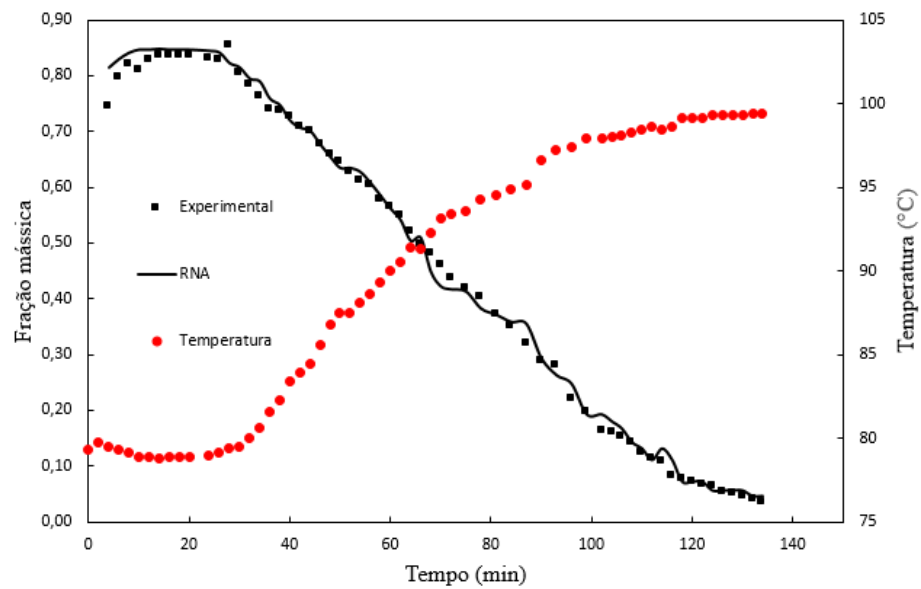


Figura 17 - Perfil de concentração do etanol durante o processo de destilação da mistura binária - Experimento 1. Concentração inicial: 12,7%(v/v).

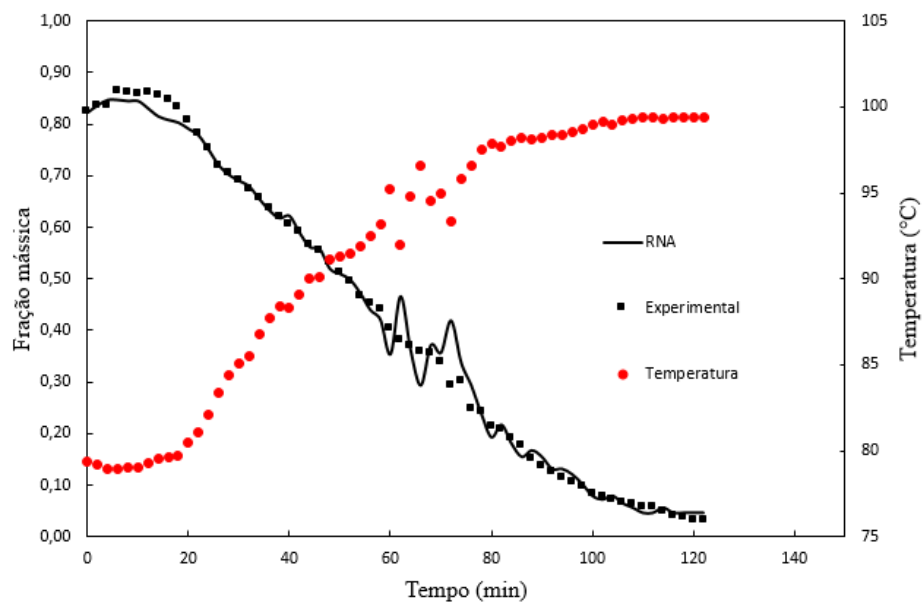


Figura 18 - Perfil de concentração do etanol durante o processo de destilação da mistura binária - Experimento 2. Concentração inicial: 13,1%(v/v).

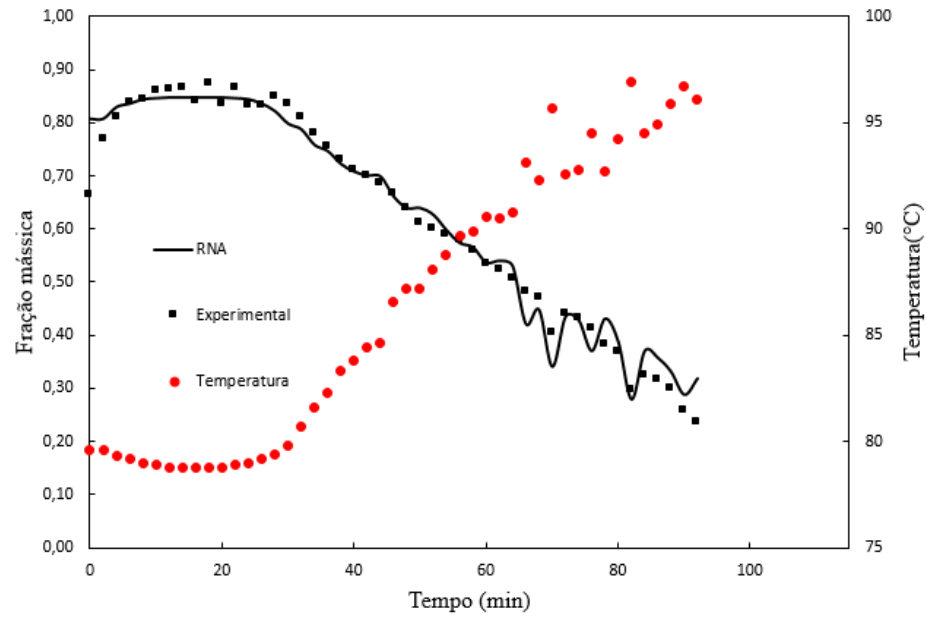


Figura 19 - Perfil de concentração do etanol durante o processo de destilação da mistura binária - Experimento 3. Concentração inicial: 10,4%(v/v).

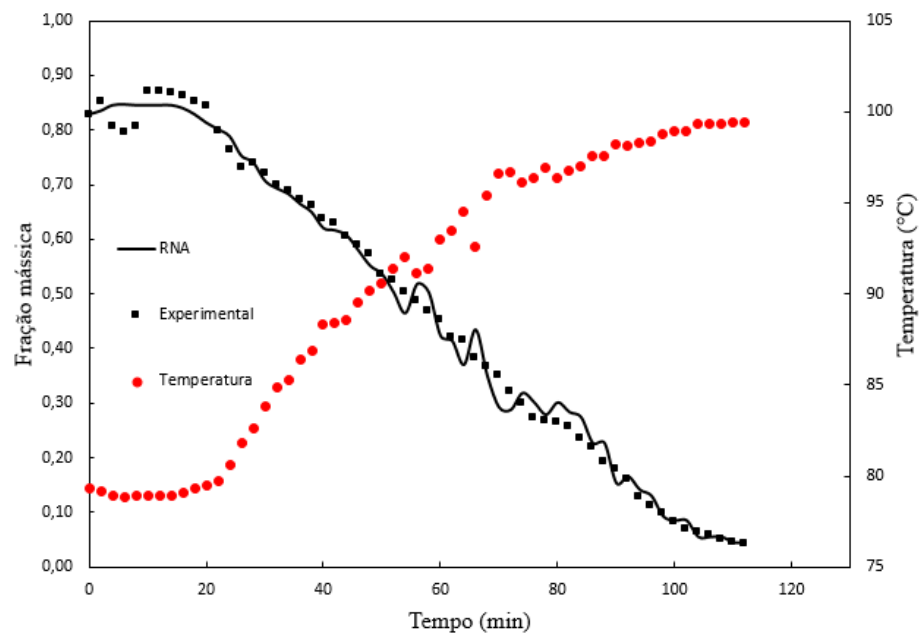


Figura 20 - Perfil de concentração do etanol durante o processo de destilação da mistura binária - Experimento 4. Concentração inicial: 12,7%(v/v).

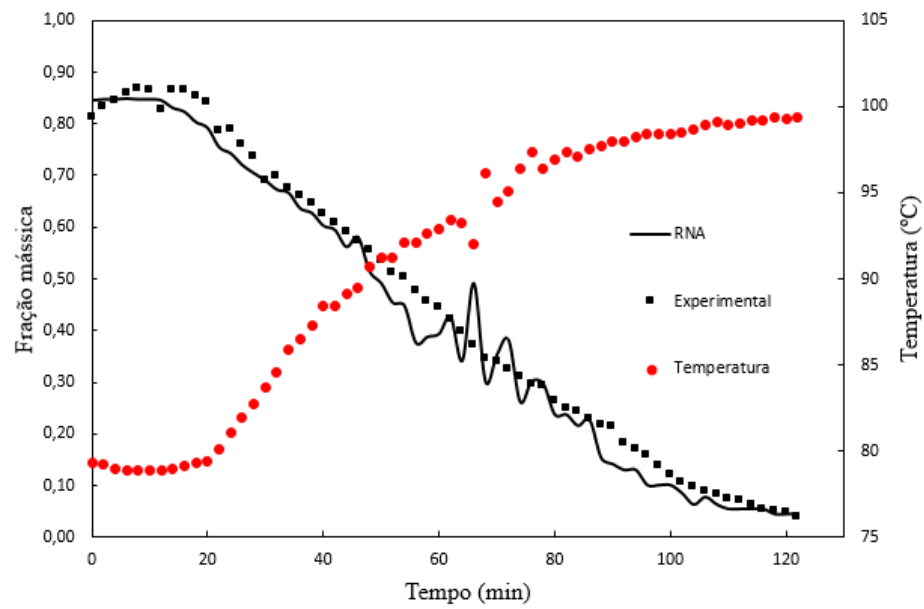


Figura 21 - Perfil de concentração do etanol durante o processo de destilação da mistura binária - Experimento 5. Concentração inicial: 13,4%(v/v).

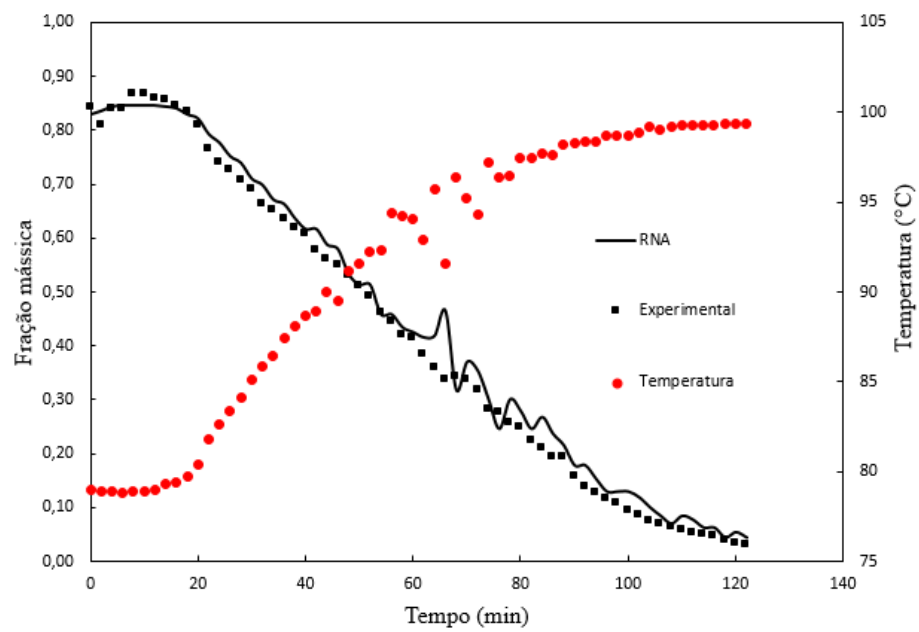


Figura 22 - Perfil de concentração do etanol durante o processo de destilação da mistura binária - Experimento 6. Concentração inicial: 12,4%(v/v).

Tabela 6 - Avaliação de desempenho do sensor virtual - Destilações das misturas sintéticas.

Experimento	EAM	Desvio padrão do erro	Erro máximo	Erro mínimo
1	0,0140	0,0126	0,0708	0,0004
2	0,0165	0,0209	0,1244	0,0003
3	0,0237	0,0253	0,1444	0,0011
4	0,0198	0,0146	0,0578	0,0004
5	0,0311	0,0240	0,1181	0,0008
6	0,0243	0,0184	0,1290	0,0019

Analisando as Figuras 17 a 22 observa-se que apesar de alguma dificuldade em obter respostas corretas para oscilações bruscas de temperatura em determinadas situações, como na Figura 21, o sensor virtual realiza ótima representação do comportamento do processo, confirmando a habilidade das Redes Neurais Artificiais no mapeamento de sistemas não lineares através de dados de entrada e saída. Através da Tabela 6 pode-se concluir que o sensor possui boa acurácia e precisão devido aos baixos valores de erro absoluto médio e desvio padrão do erro apresentados.

5.3 TESTES COM AS DESTILAÇÕES DE MOSTOS FERMENTADOS

A etapa seguinte consiste em testar a capacidade de inferência do sensor virtual proposto durante a destilação de mostos fermentados, objetivo final da sua aplicação. Neste sentido, o sensor virtual foi testado em destilações de mostos fermentados de melancia, manga e jabuticaba, cujas concentrações iniciais estão dentro da faixa que o sensor virtual foi treinado. Nas Figuras 23, 24 e 25 encontram-se os testes realizados com os dados das destilações do mosto fermentado de melancia, cuja concentração inicial foi de 11%(v/v). A Tabela 7 analisa o desempenho do sensor virtual para cada experimento.

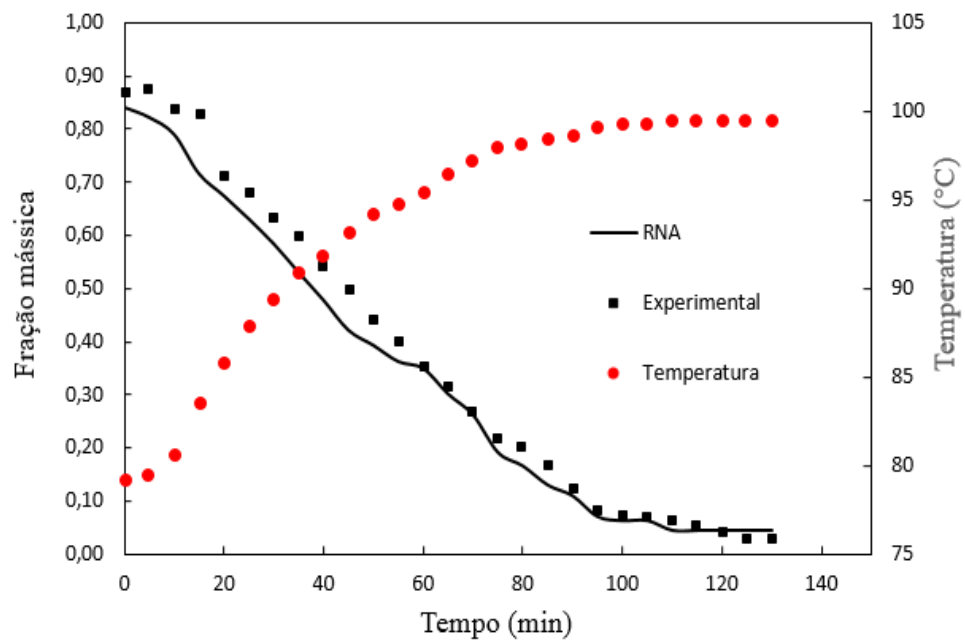


Figura 23 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de melancia - Experimento 1.

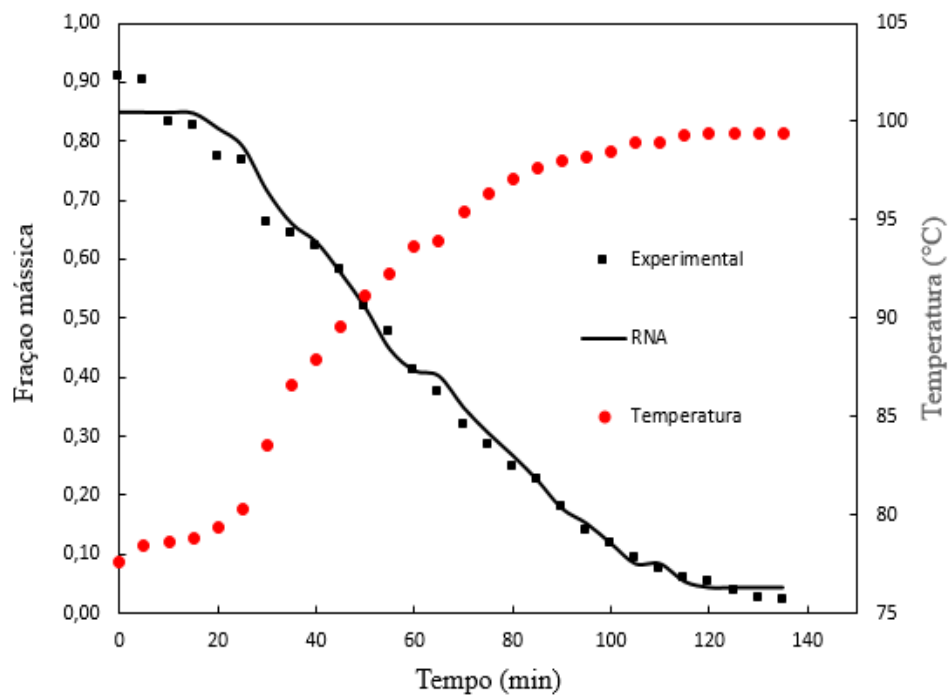


Figura 24 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de melancia - Experimento 2.

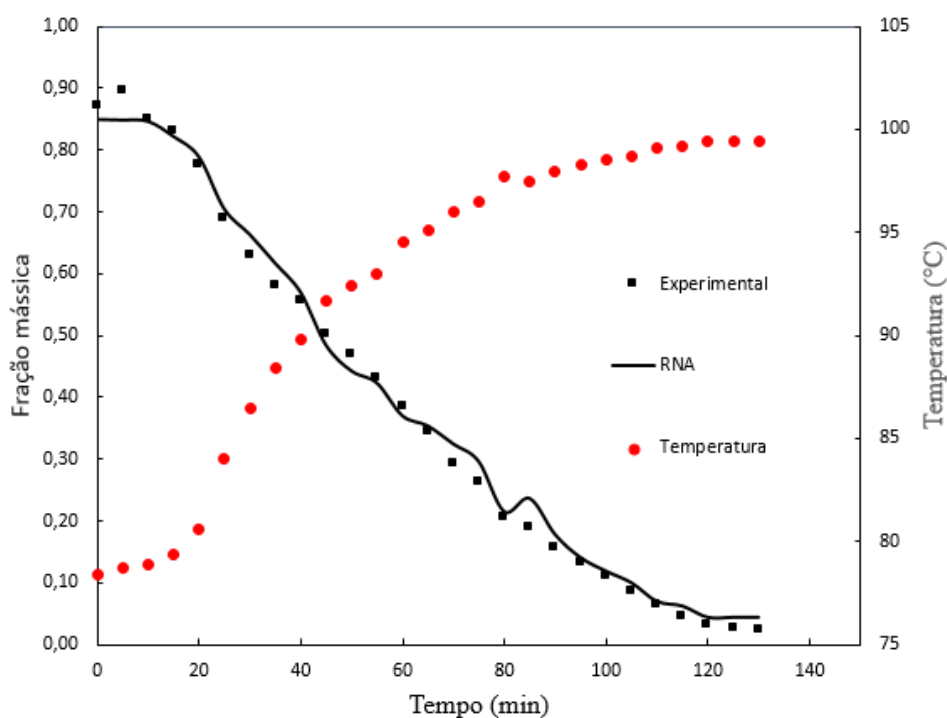


Figura 25 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de melancia - Experimento 3.

Tabela 7 - Análise de desempenho do sensor virtual - Destilações do mosto fermentado de melancia.

Experimento	EAM	Desvio padrão do erro	Erro máximo	Erro mínimo
1	0,0324	0,0261	0,1121	0,0015
2	0,0194	0,0172	0,0590	0,0006
3	0,0193	0,0119	0,0471	0,0019

Através das Figuras 23 a 25 e dos valores do EAM e do desvio padrão do erro na Tabela 7, nota-se que o sensor virtual não apresentou perda de desempenho nas destilações do mosto fermentado de melancia, tendo fornecido ótima representação do comportamento do processo. As Figuras 26 a 28 apresentam os resultados dos testes realizados com os dados das destilações do mosto fermentado de manga, de concentração inicial de 10%(v/v). A Tabela 8 analisa o desempenho do sensor para cada experimento.

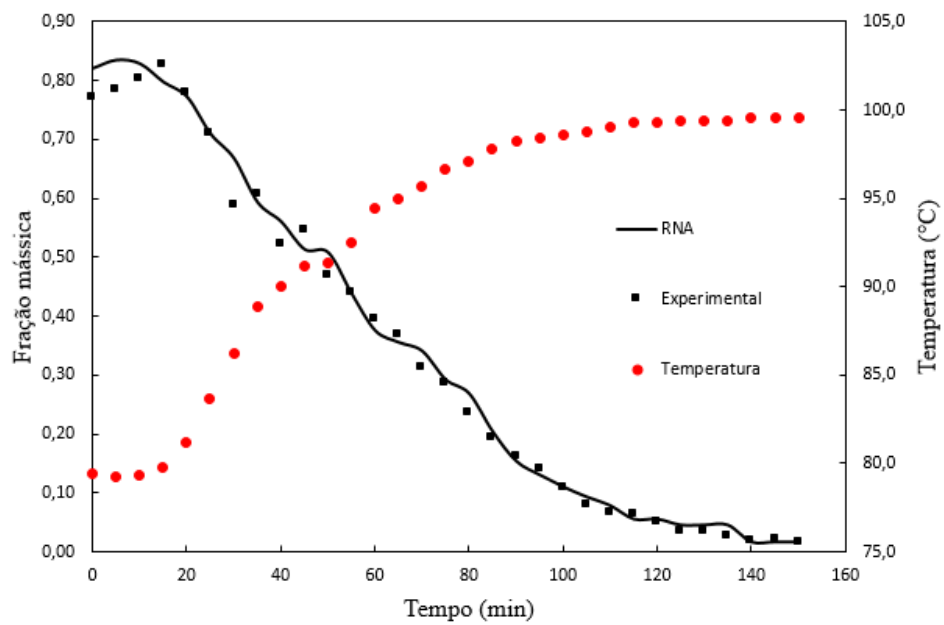


Figura 26 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de manga - Experimento 1.

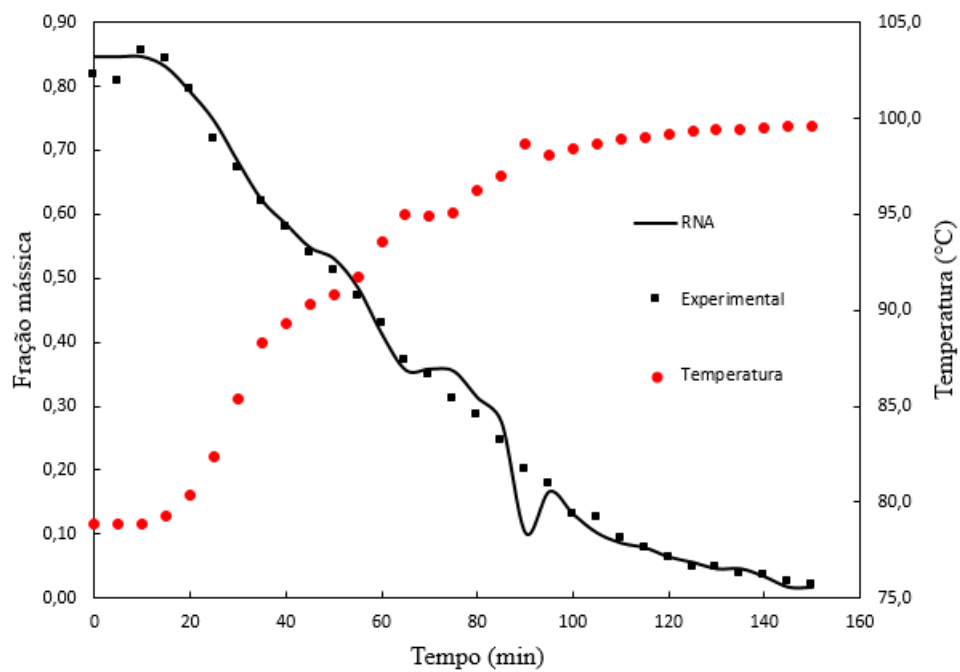


Figura 27 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de manga - Experimento 2.

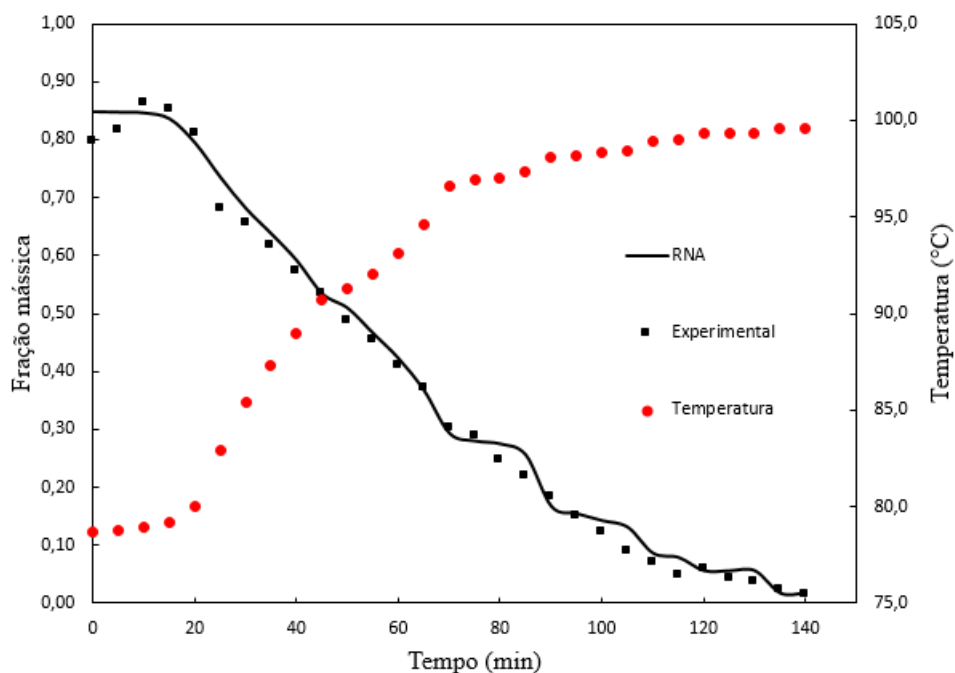


Figura 28 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de manga - Experimento 3.

Tabela 8 - Análise de desempenho do sensor virtual - Destilações do mosto fermentado de manga.

Experimento	EAM	Desvio padrão do erro	Erro máximo	Erro mínimo
1	0,0186	0,0181	0,0793	0,0001
2	0,0160	0,0192	0,1013	0,0000
3	0,0190	0,0133	0,0543	0,0013

Novamente o sensor mostrou ótimo desempenho, apresentando baixos valores de erro absoluto médio e desvio padrão do erro e tendo representado de forma satisfatória o perfil dinâmico do processo, como pode ser visto nas Figuras 26 a 28. O perfil do processo durante as destilações do mosto de jabuticaba encontra-se nas Figuras 29 a 32 e na Tabela 9 estão os resultados que dizem respeito à performance do sensor virtual.

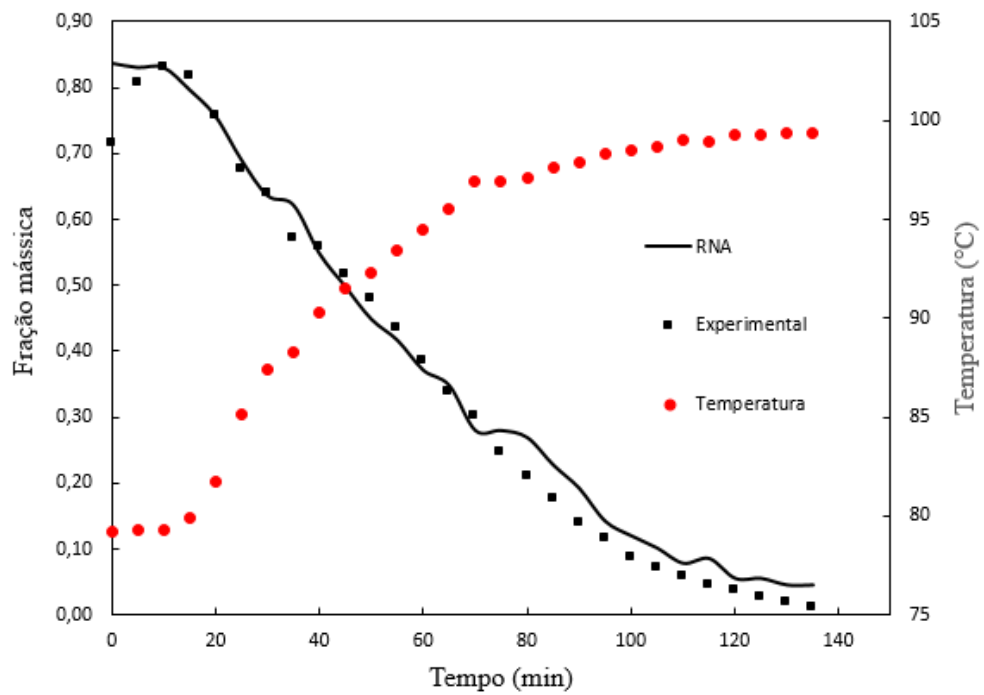


Figura 29 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de jabuticaba - Experimento 1.

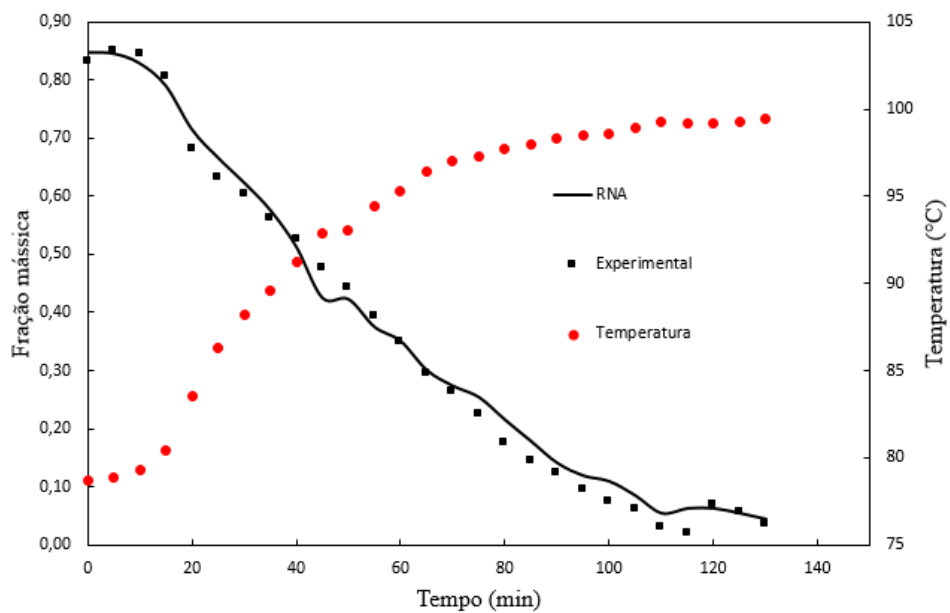


Figura 30 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de jabuticaba - Experimento 2.

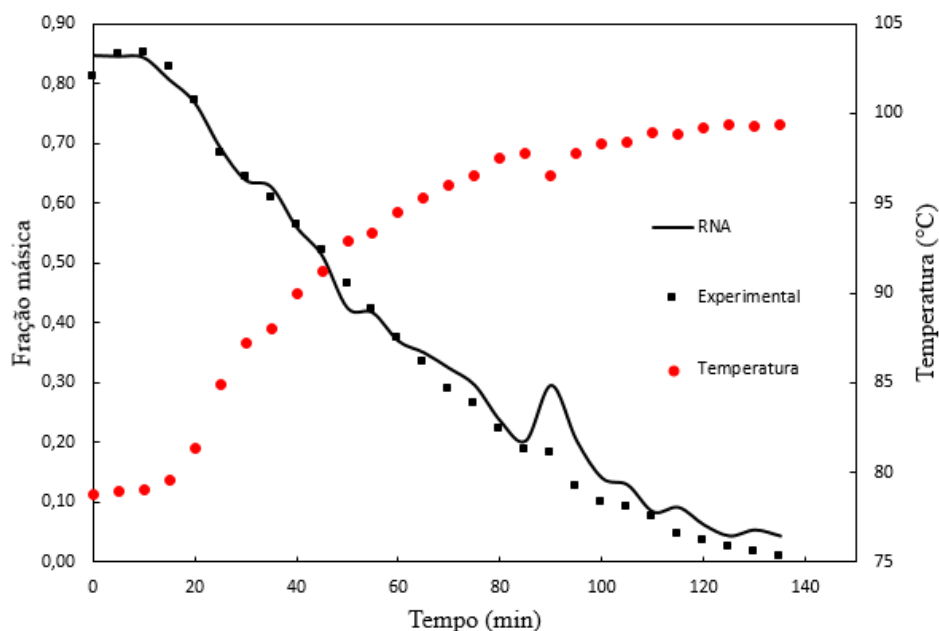


Figura 31 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de jabuticaba - Experimento 3.

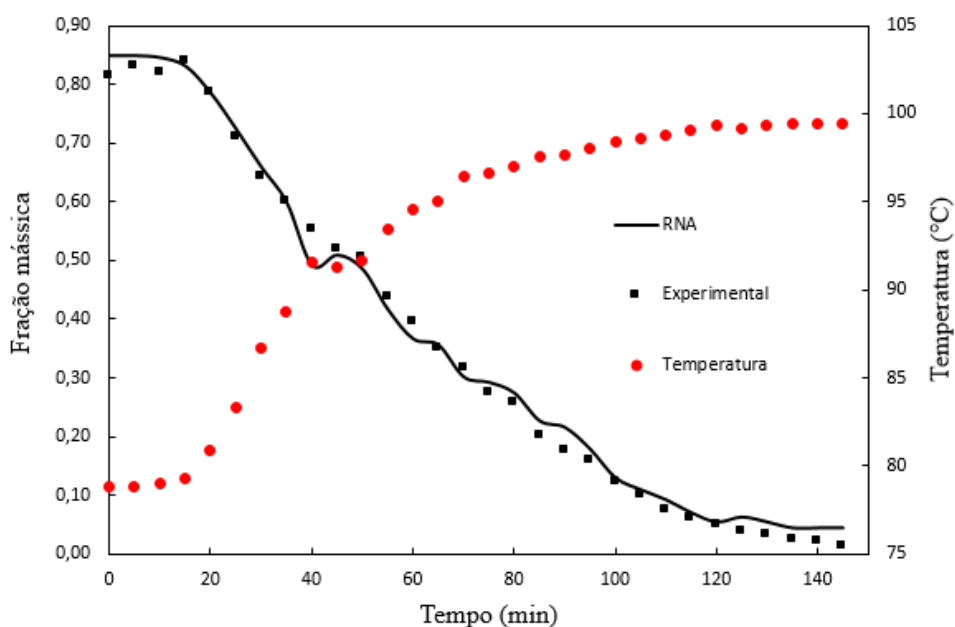


Figura 32 - Perfil de concentração do etanol durante o processo de destilação do mosto fermentado de jabuticaba - Experimento 4.

Tabela 9 - Análise de desempenho do sensor virtual - Destilações do mosto fermentado de jabuticaba.

Experimento	EAM	Desvio padrão do erro	Erro máximo	Erro mínimo
1	0,0288	0,0231	0,0121	0,0014
2	0,0208	0,0134	0,0519	0,0002
3	0,0257	0,0244	0,1145	0,0011
4	0,0191	0,0121	0,0621	0,0029

O sensor também apresentou ótimo desempenho na representação dos dados das destilações do mosto fermentado de jabuticaba, como pode ser observado nas Figuras 28 a 31 e nos resultados da Tabela 9.

Assim como nas destilações das misturas binárias, o sensor também mostrou uma queda de desempenho quando apresentado a maiores oscilações de temperatura. Tal situação pode ser observada nas Figuras 27 e 31, representada pelos picos da curva plotada. É possível que a adição de temperaturas de outros pontos da coluna como variáveis de entrada venha a minimizar o problema. Apesar disso, pode-se afirmar que o sensor obteve ótimo desempenho em todas as situações a que foi apresentado e se mostra viável para a aplicação em situações reais, sendo uma opção barata e confiável para a realização do monitoramento dos processos de destilação de bebidas fermentadas.

6 CONCLUSÃO

O sensor virtual desenvolvido apresentou ótimo desempenho na representação dos dados experimentais das destilações das misturas binárias e dos mostos fermentados, que pode ser visto pelos baixos valores de EAM e DP. Além disso, o sensor acompanhou de forma satisfatória as oscilações do processo, ratificando a capacidade das Redes Neurais em mapear funções não lineares através de conjuntos de dados de entrada e saída.

Na determinação da composição do etanol durante as destilações das misturas binárias, o sensor virtual apresentou EAM de 0,0140; 0,0165; 0,0237; 0,0198; 0,0311 e 0,0243 para cada experimento executado, respectivamente. Nos experimentos com o mosto fermentado de melancia, o sensor virtual manteve a performance apresentada na inferência das destilações binárias, com EAM igual a 0,0324; 0,0194 e 0,0193. Tal comportamento também foi observado nas destilações dos mostos de manga e jabuticaba, que apresentaram valores de EAM de 0,0186; 0,0192; 0,0190 e 0,0288; 0,0208; 0,0257 e 0,0191, respectivamente. Assim, a aproximação do comportamento dos mostos fermentados por uma mistura binária se mostrou adequada.

Foram observadas dificuldades em inferir a composição do etanol quando a temperatura do estágio de topo, utilizada como variável de entrada, apresenta instabilidade, tanto para as misturas binárias quanto para as destilações dos mostos fermentados. Entretanto, tais situações não prejudicaram o desempenho do sensor. É possível que a utilização de temperaturas de outros estágios da coluna como entrada reduzam ou eliminem esse problema. Outras arquiteturas e algoritmos de treinamento também podem ser utilizados a fim de melhorar a sua performance e eliminar essas instabilidades pontuais.

Desta forma, o sensor se mostra viável para o uso em aplicações reais, cumprindo a lacuna da falta de instrumentos baratos e eficientes para o monitoramento de processos de destilação de bebidas fermentadas e agregando tecnologia e qualidade à sua produção, principalmente para indústrias que não possuem grande capacidade de investimento em instrumentação.

REFERÊNCIAS

- ABDULLAH, Z.; AZIZ, N.; AHMAD, Z. Nonlinear modeling application in distillation column. **Chemical Product and Process Modeling**, v.2, p. 1-25, 2007.
- ALCARDE, A. R.; SOUZA, P. A. DE; BOSQUEIRO, A. C.; BELLUCO, A. E. DE S. Cinética de volatilização de componentes secundários da aguardente de cana-de-açúcar durante dupla destilação em alambique simples. **Brazilian Journal of Food Technology**, v. 13, p.271-278, 2010.
- ANSARI, H. R.; ZAREI, M. J.; SABBAGHI, S.; KESHAVARZ, P. A new comprehensive model for relative viscosity of various nanofluids using feedforward backpropagation MLP neural networks. **International Communications in Heat and Mass Transfer**, v. 91 p. 158–164, 2018.
- ARIETA-GARAY, Y.; BLANCO, P.; LÓPEZ-VÁZQUEZ, C.; RODRÍGUEZ-BENCOMO, J. J.; PÉREZ-CORREA, J. R.; LÓPEZ, F.; ORRIOLS, I. Effects of distillation system and yeast strain on the aroma profile of albariño (*vitis vinífera l.*) grape pomace spirits. **Journal of Agricultural and Food Chemistry**, v. 62, p. 10552-10560, 2014.
- ASQUIERI, EDUARDO R.; SILVA, ALINE G. DE M. E; CÂNDIDO, M. A. Aguardente de jabuticaba obtida da casca e borra da fabricação de fermentado de jabuticaba. **Ciência e Tecnologia de Alimentos**, v. 29, p. 896-904, 2009.
- BATISTA, F. R. M.; MEIRELLES, A. J. A. Computer simulation applied to studying continuous spirit distillation and product quality control. **Food Control**, v. 22, p. 1592-1603, 2011.
- BORTOLETTO, A. M.; ALCARDE, A. R. Assesment of chemical quality of brazilian sugar cane spirits and cachaças. **Food Control**, v. 54, p. 1-6, 2015.
- BRASIL. Decreto nº 6871, de 4 de junho de 2009. Dispõe sobre a padronização, a classificação, o registro, a inspeção, a produção e a fiscalização de bebidas. **Diário Oficial da União**, Brasília, DF, 4 jun. 2009.
- BRASIL. Instrução Normativa MAPA nº 15, de 01 de abril de 2011. Estabelece a complementação dos padrões de identidade e qualidade das bebidas alcoólicas destiladas em todo o território nacional. **Diário Oficial da União**, Brasília, DF, 2011.
- BRUNO, S. N. F.; VAITSMAN, D. S.; KUNIGAMI, C. N.; BRASIL, M. G. Influence of the distillation process from Rio de Janeiro in the ethyl carbamate formation in brazilian sugar cane sprits. **Food Chemistry**, v. 104, p. 1345-1352, 2007.
- CALIARI, M.; MARQUES, F. P. P.; LACERDA, D. B. C. L.; CASTRO, M. V. L. DE; SILVA, M. A. P. DA. Produção de aguardente de manda e bebida alcoólica mista de manga com diferentes fontes alcoólicas. **Revista Verde de Agroecologia e Desenvolvimento Sustentável**, v. 7, p. 175-180, 2012.

CLAUS, M. J.; BERGLUND, K. A. Defining still parameters using chemcad batch distillation model for modeling fruit spirits distillations. **Journal of Food Process Engineering**, v. 32, p. 881-892, 2009.

COSTA, D. G. **Reaproveitamento da calda da desidratação osmótica da casca de jabuticaba (*Myrciaria cauliflora*) para a produção de aguardente**. 2016. 53 f. Monografia (Bacharel em Engenharia de Alimentos) - Departamento de Tecnologia de Alimentos, Universidade Federal de Sergipe. São Cristóvão, Brasil, 2016.

COUTINHO, E. P.; RAMOS, Z. N. S.; ALVES, A. S.; OLIVEIRA, R. E. S. Boas práticas de fabricação de cachaça de alambique: visão técnica versus visão empresarial. **UNOPAR Científica Ciências Biológicas e da Saúde**, v. 14, p. 165-170, 2012.

CYBENKO, G. Approximation by Superpositions of a Sigmoidal Function. **Math. Control Signals Systems**, v. 2, p. 303-314, 1989.

DAVOUDI, E.; VAFERI, B. Applying artificial neural networks for systematic estimation of degree of fouling in heat exchangers. **Chemical Engineering Research and Design**, v. 130 p. 138–153, 2018.

DCI - DIÁRIO COMÉRCIO INDÚSTRIA E SERVIÇOS. Cachaça é opção de renda para pequenos produtores, 2017. < /www.dci.com.br> acessado em 1/03/2019.

DECLoux, M.; COUSTEL, J. Simulation of a neutral spirit production plant using beer distillation. **International Sugar Journal**, v. 107, p. 628-643, 2005.

DORAISWAMI, R.; CHEDED, L. Robust model-based soft sensor: design and application. **In: World Congress The International Federation of Automatic Control**, 19, 2014, Cidade do Cabo. Anais. Cidade do Cabo, IFAC, 2014, p. 5491-5496.

ESTEBAN-DECLoux, M.; DETERRE, S.; KADIR, S.; GIAMPAOLI, P.; ALBERT, J.; JOULIA, X.; BAUDOUIN, O. Two industrial examples of coupling experiments and simulations for increasing quality and yield of distilled beverages. **Food and Bioproducts Processing**, p. 1-12, 2014.

FORTUNA, L.; GRAZIANI, S.; RIZZO, A.; XIBILIA, M. G. **Soft Sensores for Monitoring and Control of Industrial Processes**. London: Springer, 2007.

FORTUNA, L.; GRAZIANI, S.; XIBILIA, M. G. Soft sensors for product quality monitoring in debutanizer distillation columns. **Control Engineering Practice**, v. 13, p. 499-508, 2005.

FUKUDA, T. Theory and applications of neural networks for industrial control systems. **IEE Transactions on industrial electronics**, v. 39, p.472-489, 1992.

GALANO, E.; IMBELLONI, M.; CHAMBERY, A.; ANTONIO, M., AMORESANO, A. Molecular fingerprint of the alcoholic Grappa beverage by mass spectrometry techniques. **Food Research International**, v. 75, p. 106-114, 2015.

GARCÍA-LLOBODANIN, L.; ACHAERANDIO, M. F.; GUELL, C.; LOPEZ, F. Pear distillates from pear juice concentrate: effect of lees in the aromatic composition. **Journal of Agricultural and Food Chemistry**, v. 55, p. 3462-3468, 2007.

GAVIN, H. P. The Levenberg-Marquardt method for nonlinear least squares curve fitting problems. **Department of Civil and Environmental Engineering -Duke University**, 2017.

GRBIC, R.; SLISKOVIC, D.; KADLEC, P. Adaptive soft sensor for online prediction and process monitoring based on a mixture of Gaussian process models. **Computers and Chemical Engineering**, v.58, p. 84-97, 2013.

GURNEY, K. **An introduction to neural networks**. Londres: CRC Press, 1997.

HAGAN, M. T.; DEMUTH, H. B.; BEALE, M. H.; DE JESÚS, O.. **Neural Network Design**. 2ª Edição, 2014.

HAGAN, M. T.; MENHAJ, M. B. Training feedforward networks with the Marquardt algorithm. **IEE Transactions on Neural Networks**, v. 5, p. 989-993, 1994.

HASSANPOUR, M.; VAFERI, B.; MASOUMI, M. E. Estimation of pool boiling heat transfer coefficient of alumina water-based nanofluids by various artificial intelligence (AI) approaches. **Applied Thermal Engineering**, v. 128, p. 1208-1222, 2018.

HAYKIN, S. **Redes Neurais: princípios e práticas**. Traduzido por Paulo Martins Engel. 2ª ed. Porto Alegre: Bookman, 2001.

HENSON, M. A. Nonlinear model predictive control: current status and future directions. **Computers and Chemical Engineering**, v. 23, p. 187-202, 1998.

HERNÁNDEZ-GÓMEZ, L. F.; ÚBEDA, J.; BRIONES, A. Melon fruit distillates: comparison of different distillation methods. **Food Chemistry**, v.82, p. 539-543, 2003.

HUANG, G.; SONG, S.; GUPTA, J. N. D.; WU, C. Semi-supervised and unsupervised extreme learning machines. **IEEE Transactions on Cybernetics**, v. 44, p. 2405-2417, 2014.

ICMC-USP: INSTITUTO DE CIÊNCIAS MATEMÁTICAS E DE COMPUTAÇÃO DA UNIVERSIDADE DE SÃO PAULO, disponível em <<http://conteudo.icmc.usp.br/pessoas/andre/research/neural/index.htm#topicos>>, consultado em 30/01/2019.

ITO, E. H. C.; SECCHI, A. R.; GOMES, M. V. C.; PAIVA, C. R. Development of a gas composition soft sensor for distillation columns: A simplified model based and robust approach. **In: International Symposium on Process Systems Engineering**, 13, 2018, San Diego. Anais. San Diego, Elsevier, 2018.

JALEE, E. A.; APARNA, K. Neuro-fuzzy Soft sensor estimator for benzene toluene distillation column. **Procedia Technology**, v. 25, p. 92-99, 2016.

JAMES, S. C.; LEGGE, R. L.; BUDMAN, H. On-line estimation in bioreactors: A review. **Reviews in Chemical Engineering**, v. 16, n. 4, p. 311-340, 2000.

JANA, A. K.; BANERJEE, S. Neuro Estimator-based Inferential Extended Generic Model Control of a Reactive Distillation Column. **Chemical Engineering Research and Design**, v. 130, p. 284-294, 2018.

KADLEC, P. Soft sensors: software-based sensors. In: WEBSTER, J. G.; EREN, H. **Measurement, instrumentation and sensors handbook**. 2 ed. Boca Raton: CRC Press, 2014. cap. 1, p. 1-13.

KADLEC, P.; GABRYS, B.; STRANDT, S. Data-driven soft sensors in the process industry. **Computers and Chemical Engineering**, v. 33, p. 795-814, 2009.

KANO, M.; MIYAZAKI, K.; HASEBE, S.; HASHIMOTO, I. Inferential control system of distillation compositions using dynamic partial least square regression. **Journal of Process Control**, v. 10, p.157-166, 2000.

KATARIA, G.; SINGH, K.; DOHARE, R.; ANN based soft sensor model for reactive distillation column. **International Journal of Advanced Technology and Engineering Exploration**, v. 3, p. 182-186, 2016.

KHOSROWSHAHI, F., Innovation in artificial neural network learning: Learn-on-demand methodology. **Automation in Construction**, v. 20, p. 1204–1210, 2011.

KRENKER, A.; BESTER, J.; KOS, A. Introduction to the artificial neural networks. In: KENJI, SUZUKI. **Artificial neural networks – Methodological advances and biomedical applications**. Rijeka, IntechOpen, 2011. cap. 1, p. 3-18.

KRÖSE, B. J.; VAN DER SMAGT, P. **An Introduction to Neural Networks**. Amsterdam: University of Amsterdam, 1996.

LÉAUTÉ, R. Distillation in Alambic. **American Society for Enology and Viticulture**, v. 41, p. 90-103, 1990.

LI, G.; ZHONG, Q.; WANG, D.; ZHANG, X.; GAO, H.; SHIGANG, S. Determination and formation of ethyl carbamate in chinese spirits. **Food Control**, v. 56, p. 169-176, 2015.

LIMA, U. DE A.; TEIXEIRA, C. G.; BERTOZZI, J. C.; SERAFIM, F. A. T.; ALCARDE, A. R. Influence of fast and slow distillation on ethyl carbamate content and on coefficient on non-alcohol components in brazilian sugarcane spirits. **Institute of Brewing and Distilling**, v. 118, p. 305-308, 2012.

LOLLI, F.; GAMBERINI, R.; REGATTIERI, A.; BALUGANI, E.; GATOS, T.; GUCCI, S. Single-hidden layer neural networks for forecasting intermittent demand. **International Journal Production Economics**, v. 183 p. 116–128, 2017.

LOTUFO, F. A.; GARCIA, C. Sensores virtuais ou soft sensors: uma introdução. In: **Brazilian Conference on Dynamics, Control and Applications**, 7, 2008, Presidente Prudente. Anais. Presidente Prudente, FCT, 2008, p. 1-9.

LUCAS, M. **Sensor virtual inteligente para estimação de composições em colunas de destilação**. 2012. 230 f. Dissertação (Mestrado em Engenharia Elétrica) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2012.

LUNA, R.; LÓPEZ, F.; PÉREZ-CORREA, J. R. Minimizing methanol content in experimental charentais alembic distillations. **Journal of Industrial and Engineering Chemistry**, v. 57 p. 160–170, 2018.

MADRERA, R. R.; GOMIS, D. B.; ALONSO, J. J. M. Influence of distillation system, oak wood type, and aging time on volatile compounds of cider brandy. **Journal of Agricultural and Food Chemistry**, v. 51, p. 5709-5714, 2003.

MEHTA, S.; RAMANI, H; NILESHKUMAR, N.; RAHMAN, Y.; RAHMAN, I. Recursive orthogonal least square based soft sensor for batch distillation. **Chemical Product and Process Modeling**, v. 12, n. 3, p. 241-263, 2016.

MEIRELES, M. R. G.; ALMEIDA, P. E. M. A comprehensive review for industrial applicability of artificial neural networks. **IEE Transactions on industrial electronics**, v. 50, p. 585-601, 2003.

MOHLER, I.; ANDRIJIC, Z. U.; BOLF, N. Development of soft sensors for crude distillation unit control. **In: IFAC World Congress**, 18, 2011, Milão. Anais. Milão, IFAC, 2011.

MOTA, L. O. **Predição e controle da qualidade no processo de destilação em batelada com o uso de sensor virtual**. 2017. 61 f. Dissertação (Mestrado em Engenharia Química). Universidade Federal de Sergipe, Programa de Pós Graduação em Engenharia Química, São Cristóvão, Brasil, 2017.

OECD - ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. Brazil joins the OECD Fruits and Vegetables Scheme, 2018. <<http://www.oecd.org/agriculture/fruit-vegetables/>>, acessado em 30/01/2019.

OISIOVICI, R. M.; CRUZ, S. L. State estimation of batch distillation columns using an extended Kalman filter. **Chemical Engineering Science**, v. 55, p. 4667-4680, 2000.

OLIVEIRA, P. N. DE; GOMES, P. C. S.; SANTOS, D. E. L.; ALCARDE, A. R.; ABUD, A. K. DE S.; OLIVEIRA JUNIOR, A. M. DE O. Produção de aguardente de manga (*Mangifera indica* L. cv Tommy Atkins) para valorização da agroindústria tropical. **III Congresso Internacional de Gastronomia e Ciência de Alimentos**, 2018.

OSORIO, D.; PÉREZ-CORREA, J. R.; AGOSINA, E.; CABRERA, M. Soft-sensor for online estimation of ethanol concentrations in wine stills. **Journal of Food Engineering**, v. 87, n. 4, p. 571-577, 2008.

OSORIO, D.; PÉREZ-CORREA, J. R.; BIEGLER, L. T., AGOSIN, E. Wine distillates: practical operating recipe formulation for stills. **Journal of Agricultural and Food Chemistry**, v. 53, p. 6326-6331, 2005.

PANI, A. K.; MOHANTA, H. K. A survey of data treatment techniques for soft sensor design. **Chemical Product and Process Modeling**, v. 6, p. 1-23, 2011.

PARENTE, G. D. L. **Cinética da fermentação e da destilação na produção de aguardente de abacaxi**. 2014, 77 f. Dissertação (Mestrado em Sistemas Agroindustriais). Centro de Ciências e Tecnologia Agroalimentar - Universidade Federal de Campina Grande. Pombal, Brasil, 2014.

PATIL, S. R.; GHATE, V. N. A generalized regression neural network based on soft sensor for multicomponent distillation column. **International Journal of Computer Communication Engineering**, v. 4, p. 371-378, 2015.

PECIC, S.; VELJOVIC M., DESPOTOVIC, S.; LESKOSEK, I.; JADRANIN, M.; TESEVIC, V.; NIKSIC, M.; NIKSEVIC, N. Effect of maturation conditions on sensory and antioxidant properties of old serbian plum brandies. **European Food Research and Technology**, v. 235, p. 479-487, 2012.

PUNTES, C.; JOULIA, X.; VIDAL, J.; ESTEBAN-DECLoux, M. Simulation of spirits distillation for a better understanding of volatile aroma compounds behavior: application to armagnac production. **Food and Bioproducts Processing**, v. 112, p. 31-62, 2018.

RAMLI, M. N.; HUSSAIN, M. A.; JAN, B. M.; ABDULLAH, B. Composition prediction of a debutanizer column using equation based artificial neural network model. **Neurocomputing**, v. 131, p. 59-76, 2014.

RANI, A.; SINGH, V.; GUPTA, J.R.P. Development of soft sensor for neural network based control of distillation column. **ISA Transactions**, v. 52, p. 438-449, 2013.

ROGINA, A.; SISKI, I.; MOHLER, I.; UJVIC, Z.; BOLF, N. Soft sensor for continuous product quality estimation (in crude distillation unit). **Chemical Engineering Research and Design**, v. 89, p. 2070-2077, 2011.

ROJAS, R. **Neural Networks: A systematic introduction**. Berlin: Springer-Verlag, 1996.

RUSSEL, S.; NORVIG, P. **Artificial Intelligence: A modern approach**. 3ª Edição. Upper Saddle River: Prentice Hall, 2003.

SACHER, J.; GARCIA-LLOBODANIN, L.; LOPEZ, F.; SEGURA, H.; PÉREZ-CORREA, J. R. Dynamic modeling and simulation of na alembic pear wine distillation. **Food and Bioproducts Processing**, v. 91, p. 447-456, 2013.

SCAVANINI, H. F.A; CERIANI, R.; MEIRELLES, A. J. *Cachaça* distillation investigated on the basis of model systems. **Brazilian Journal of Chemical Engineering**, v. 29, p. 429-440, 2012.

SHAO, W.; TIAN, X. Adaptive soft sensor for quality prediction of chemical processes based on selective ensemble of local partial least squares models. **Chemical Engineering Research Design**, v. 95, p.113-132, 2015.

SHI, X.; XIONG, W. Approximate linear dependence criteria with active learning for smart soft sensor design. **Chemometrics and Intelligent Laboratory Systems**, v.180, p. 88-95, 2018.

SILVA, I. N. DA; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais para engenharia e ciências aplicadas: curso prático**. 2ª Edição. São Paulo: Altiber, 2010.

SILVA, IVAN N.; SPATTI, D. H.; FLAUZINO, R. A.; LIBONI, L. H. B.; ALVES, SILAS F. R. **Artificial neural networks: A practical course**. Springer, 2016.

SILVA, L. N. DE C. **Análise e síntese de estratégias de aprendizado para redes neurais artificiais**. 1998. 248 f. Dissertação (Mestrado em Engenharia de Computação e Automação Industrial). Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, Campinas, Brasil, 1998.

SILVA, M. B. DE L. DA; C., J. B. P.; LELIS, V. G., ALVARENGA, L., M.; ZUIM, D. R., SILVA, P. H. A. DA. Qualidade físico-química e sensorial de aguardentes de polpa de banana e banana integral submetidas à hidrólise enzimática. **Alimentos e Nutrição**, v. 20, p. 217-221, 2009.

SLISKOVIC, D.; GRBIC, R.; HOCENSKI, Z. Methods for plant data-based process modeling in soft-sensor development. **Automatika**, v. 52, p. 306-318, 2011.

SMETS, I. Y.; BOON, S.; BOELEN, T.; ESPINOSA, J.; IMPE, J. F. VAN. Inferring distillation product composition: a hybrid soft sensor approach. In: **Symposium on Dynamics and Control of Process Systems**, 8, 2007, Cancún. Anais. Cancún, v. 3, IFAC, 2007, p. 167-172.

SOUZA, F. A. A.; ARAÚJO, R.; MENDES, J. Review of soft sensor methods for regression applications. **Chemometrics and Intelligent Laboratory Systems**, v. 152, p. 69-79, 2016.

SPAHO, N. Distillation techniques in the fruit spirits production. In: MENDES, MARISA. **Distillation – Innovative applications and modeling**. Londres, IntechOpen, 2017. Cap. 6, p. 129-152.

SPAHO, N.; DURR, P.; GRBA, S.; VELAGIC-HABUL, E.; BLESIC, M. Effects of distillation cut on the distribution of higher alcohols and esters in brandy produced from three plum varieties. **Institute of Brewing and Distilling**, v. 119, p. 48-56, 2013.

SUN, K.; HUANG, SH.; JANG, S. S.; WONG, D. SH. Development of soft sensor with neural network and nonlinear variable selection for crude distillation unit process. **Computer Aided Chemical Engineering**, v. 38, p. 337-342, 2016.

TORGASHOV, A.; ZMEU, K. Identification of nonlinear soft sensor models of industrial distillation process under uncertainty. **IFAC-Papers Online**, v. 48, n. 28, p. 45-46, 2015.

VON ZUBEN, F. J. **Modelos paramétricos e não-paramétricos de redes neurais artificiais e aplicações**. 1996. 244 f. Tese (Doutorado em Engenharia Elétrica). Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica, Campinas, Brasil, 1996.

YELGATTE, N. N.; RAUT, S. A. MEHTA, S.; RAHMAN, I. OLS based-RBF Networks for Estimation of Product Composition in Batch Distillation. **International Journal of Chemical Sciences and Applications**, v. 6, p. 64-71, 2015.

ZAMPROGNA, E.; BAROLO, M.; SEBORG, D. E. Development of a soft sensor for a batch distillation column using linear and nonlinear PLS regression techniques. **In: Triennial World Congress**, 15, 2002, Barcelona. Anais, Barcelona, IFAC, 2002, p. 431-436.

ZHANG, S.; BO, C.; SUN, C.; WANG, Y. Study on the soft sensor and control scheme for na industrial azeotropic distillation column. **Advanced Control of Chemical Process**, p. 137-142, 2003.