

UNIVERSIDADE FEDERAL DO MARANHÃO  
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE ELETRICIDADE

**RAQUEL MACHADO DE SOUSA**

*Avaliação da Predição de Algoritmos de Treinamento  
Supervisionado de Redes Neurais Artificiais Aplicado a  
Qualidade de Biodiesel*

São Luís  
2015

RAQUEL MACHADO DE SOUSA

*Avaliação da Predição de Algoritmos de Treinamento  
Supervisionado de Redes Neurais Artificiais Aplicado a  
Qualidade de Biodiesel*

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Eletricidade da Universidade Federal do Maranhão, para a obtenção do título de mestre em Engenharia de Eletricidade - Área de Concentração: Ciência da Computação.

**Orientador: Dr. Sofiane Labidi**

São Luís

2015

Sousa, Raquel Machado de.

Avaliação da predição de algoritmos de treinamento supervisionado de redes neurais artificiais aplicado a qualidade de biodiesel / Raquel Machado de Sousa. – São Luís, 2015.

80 f.

Impresso por computador (fotocópia).

Orientador: Sofiane Labidi.

Dissertação (Mestrado) – Universidade Federal do Maranhão, Programa de Pós-Graduação em Engenharia de Eletricidade, 2015.

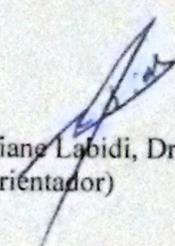
1. Biodiesel. 2. Redes neurais artificiais. 3. Índice de iodo. 4. Predição. 5. Viscosidade. I. Título.

CDU 662.756.3

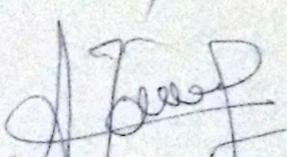
**AVALIAÇÃO DA PREDIÇÃO DE ALGORITMOS DE  
TREINAMENTO SUPERVISIONADO DE REDES NEURAIAS  
ARTIFICIAIS APLICADO A QUALIDADE DE BIODIESEL**

**Raquel Machado de Sousa**

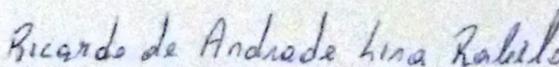
Dissertação aprovada em 26 de fevereiro de 2015.



Prof. Sofiane Labidi, Dr.  
(Orientador)



Prof. Zair Abdelouahab, Ph.D.  
(Membro da Banca Examinadora)



Prof. Ricardo de Andrade Lira Rabêlo, Dr.  
(Membro da Banca Examinadora)

*Dedico a minha avó Beatriz,  
pelo amor incondicional.*

## RESUMO

Para assegurar uma combustão eficiente e qualidade nas emissões, bem como a segurança no transporte e manuseio do biodiesel, a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) instituiu, por meio da Resolução nº 14 de 2012, parâmetros de qualidade e especificações para esse biocombustível, e para isso diversos métodos oficiais ou alternativos podem ser utilizados. Na literatura, é possível identificar uma crescente utilização de métodos lineares e não-lineares no reconhecimento e classificação de padrões aplicados ao monitoramento da qualidade de biodiesel. Nesse contexto, as Redes Neurais Artificiais (RNAs) vêm se mostrando bastante viáveis, como ferramenta não-linear, na predição de propriedades de biocombustíveis. O presente trabalho propõe avaliar a predição de propriedades de qualidade de biodiesel utilizando algoritmos de treinamento supervisionado de RNAs. De modo a contribuir com um estudo para se obter uma estrutura de rede com um algoritmo de treinamento que consiga bons resultados com melhor desempenho na predição. Por meio da predição das propriedades do biodiesel a partir da composição dos ésteres da matéria-prima, será possível avaliar a viabilidade da utilização de tal matéria-prima para a síntese de um biodiesel de qualidade. No presente trabalho obteve-se uma melhor arquitetura de RNA para predição de índice de iodo e viscosidade. Os resultados obtidos das simulações mostraram que as RNAs são uma tecnologia que pode ser utilizada para prever essas propriedades, como outras relacionadas a composição de ésteres de ácidos graxos.

**Palavras-chave:** Biodiesel, Índice de Iodo, Predição, Redes Neurais Artificiais, Viscosidade

## ABSTRACT

To ensure efficient combustion and emissions quality, as well as safety in the transport and handling of biodiesel, the National Agency of Petroleum, Natural Gas and Biofuels (ANP) establishing, through Resolution No. 14 of 2012, quality standards and specifications for this biofuel, and for that many official or alternative methods may be used. In literature, it is possible to identify an increasing use of linear methods and non - linear in the recognition and classification standards applied to the monitoring of biodiesel quality. In this context, the Artificial Neural Networks (ANN) have shown to be quite viable, as a tool non - linear, in predicting biofuel properties. The present work proposes to assess the prediction of biodiesel quality properties using supervised training algorithms of ANNs. In order to contribute to a study to provide a network structure with a training algorithm that can perform better with good results in the prediction. Through the prediction of the properties of the biodiesel from the composition of the esters of the raw material, it is possible to assess the feasibility of using such raw materials for the synthesis of a quality biodiesel. In this work we obtained a better ANN architecture for iodine value prediction and viscosity. The results of the simulations showed that the ANNs are a technology that can be used to predict these properties, like other related composition of fatty acid esters.

**Keywords:** Artificial Neural Networks, Biodiesel, Iodine Value, Prediction, Viscosity

## Agradecimentos

Agradeço a Deus, pelo seu imenso amor por mim, por ter me dado o dom da vida, por ser a fonte de força nos momentos difíceis, e por ser presença fiel nos momentos felizes da minha história. A toda minha família: aos meus pais e aos meus irmãos pelo apoio; meus tios e primos que me ajudaram muito para que eu pudesse encarar o desafio do mestrado; e, os meus padrinhos que sempre torceram pelo meu sucesso.

Aos meus melhores amigos, que mesmo estando longe fisicamente se mostraram presentes me incentivando a não desistir e ajudando sempre que eu precisava e quando a saudade apertava, especialmente, Jailson, Ismael, Raira, Myria, Daniele, Leyde, Andiará e Janaina. Eu seria menos feliz sem a presença de vocês.

Aos amigos que fiz ao longo dessa jornada em São Luis, aos amigos da Aiesec e aos amigos da UFMA que compartilhei muitos momentos divertidos e desafiadores: Thiago, Rayane, Paulo, Suzane, Dhully, Filipe, Luiz Aurélio, Nadson, Sidney, Antonio e colegas dos demais laboratórios da UFMA que tive a oportunidade de conhecer, aprendendo um pouco mais e compartilhando conhecimento.

Aos mentores e colegas que tornaram possível essa dissertação, e me mostraram o caminho a seguir numa área totalmente nova para mim, em especial, Alex Barradas, Isabelle Viegas e todos os membros dos laboratório do LAPQAP, LPQA e coordenação do PRH 39 que contribuíram para conclusão deste trabalho.

Aos professores que me encaminharam, contribuíram de alguma forma para este momento e me motivaram a querer buscar o conhecimento na minha vida escolar e acadêmica. Principalmente aos professores do mestrado, Zair Abdelouahab, Rosario Girardi, João Viana, Allan Kardec, Ewaldo Eder e demais professores que colaboraram com meu conhecimento.

Sempre serei grata em especial ao meu orientador, Sofiane Labidi, que mais do que os ensinamentos e a oportunidade do mestrado, tornou possível a conclusão desse trabalho e de mais uma etapa de minha vida.

*"Todo o homem pode, se assim o desejar, tornar-se escultor do seu próprio cérebro."*

*Santiago Ramón y Cajal*

## Lista de Figuras

2.1	Esquema da reação de transesterificação de um triacilglicerídeo catalisada por base. Fonte: [16]	22
2.2	Estruturas de ácidos graxos mais comuns presentes em óleos vegetais. Fonte: [16]	25
2.3	Modelo da célula neural biológica. Fonte: [47]	28
2.4	Modelo da célula neural artificial. Fonte: [15]	28
2.5	Principais funções de ativação. Fonte: [15]	30
2.6	Rede <i>feedforward</i> de camada simples. Fonte: [49]	31
2.7	Rede <i>feedforward</i> de camada múltipla. Fonte: [49]	31
2.8	Rede recorrente. Fonte: [49]	32
3.1	Quantidade de ésteres predominantes nas amostras de biodiesel	45
4.1	Gráfico do coeficiente de correlação do índice de iodo para backpropagation padrão com uma camada.	54
4.2	Gráfico do coeficiente de correlação do índice de iodo para backpropagation com <i>momentum</i> com uma camada.	54
4.3	Gráfico do coeficiente de correlação do índice de iodo para BFGS com uma camada	55
4.4	Gráfico do coeficiente de correlação do índice de iodo para gradiente conjugado escalonado com uma camada.	55
4.5	Gráfico do coeficiente de correlação do índice de iodo para Resilient-Propagation com uma camada.	56
4.6	Gráfico do coeficiente de correlação do índice de iodo para One-Step-Secant com uma camada.	56

4.7	Gráfico do coeficiente de correlação do índice de iodo para Levenberg-Marquardt com uma camada. . . . .	57
4.8	Gráfico do coeficiente de correlação do índice de iodo para backpropagation padrão com duas camadas. . . . .	58
4.9	Gráfico do coeficiente de correlação do índice de iodo para backpropagation com <i>momentum</i> com duas camadas. . . . .	59
4.10	Gráfico do coeficiente de correlação do índice de iodo para BFGS com duas camadas. . . . .	59
4.11	Gráfico do coeficiente de correlação do índice de iodo para gradiente conjugado escalonado com duas camadas. . . . .	60
4.12	Gráfico do coeficiente de correlação do índice de iodo para Resilient-Propagation com duas camadas. . . . .	60
4.13	Gráfico do coeficiente de correlação do índice de iodo para One-Step-Secant com duas camadas. . . . .	61
4.14	Gráfico do coeficiente de correlação do índice de iodo para Levenberg-Marquardt com duas camadas. . . . .	61
4.15	Gráfico do coeficiente de correlação da viscosidade para backpropagation padrão com uma camada. . . . .	63
4.16	Gráfico do coeficiente de correlação da viscosidade backpropagation com <i>momentum</i> com uma camada. . . . .	63
4.17	Gráfico do coeficiente de correlação da viscosidade para BFGS com uma camada. . . . .	64
4.18	Gráfico do coeficiente de correlação da viscosidade para gradiente conjugado escalonado com uma camada. . . . .	64
4.19	Gráfico do coeficiente de correlação da viscosidade para Resilient-Propagation com uma camada. . . . .	65
4.20	Gráfico do coeficiente de correlação da viscosidade para One-Step-Secant com uma camada. . . . .	65
4.21	Gráfico do coeficiente de correlação da viscosidade para Levenberg-Marquardt com uma camada. . . . .	66

4.22	Gráfico do coeficiente de correlação da viscosidade para backpropagation padrão com duas camadas. . . . .	67
4.23	Gráfico do coeficiente de correlação da viscosidade backpropagation com <i>momentum</i> com duas camadas. . . . .	68
4.24	Gráfico do coeficiente de correlação da viscosidade para BFGS com duas camadas. . . . .	68
4.25	Gráfico do coeficiente de correlação da viscosidade para gradiente conjugado escalonado com duas camadas. . . . .	69
4.26	Gráfico do coeficiente de correlação da viscosidade para Resilient-Propagation com duas camadas. . . . .	69
4.27	Gráfico do coeficiente de correlação da viscosidade para One-Step-Secant com duas camadas. . . . .	70
4.28	Gráfico do coeficiente de correlação da viscosidade para Levenberg-Marquardt com duas camadas. . . . .	70

## Lista de Tabelas

2.1	Modelos desenvolvidos para predição de propriedades de qualidade de combustíveis e biocombustíveis. . . . .	40
3.1	Configuração das topologias de redes simuladas. . . . .	49
4.1	Resultados dos testes para escolha do valor de parada antecipada para índice de iodo. . . . .	51
4.2	Resultados dos testes para escolha do valor de parada antecipada para viscosidade. . . . .	52
4.3	Coeficientes de predição para índice de iodo com redes de uma camada oculta. . . . .	53
4.4	Coeficientes de predição para índice de iodo com redes de duas camadas ocultas. . . . .	58
4.5	Coeficientes de predição para viscosidade com redes de uma camada oculta. . . . .	62
4.6	Coeficientes de predição para viscosidade com redes duas camadas ocultas. . . . .	67

# Sumário

<b>Lista de Figuras</b>	<b>viii</b>
<b>Lista de Tabelas</b>	<b>xi</b>
<b>1 Introdução</b>	<b>14</b>
1.1 Contexto . . . . .	14
1.2 Problemática . . . . .	15
1.3 Solução Proposta . . . . .	16
1.4 Objetivos . . . . .	16
1.4.1 Objetivos Específicos . . . . .	17
1.5 Metodologia de pesquisa . . . . .	17
1.6 Apresentação da Dissertação . . . . .	18
<b>2 Fundamentação Teórica</b>	<b>19</b>
2.1 Biodiesel . . . . .	19
2.1.1 Biodiesel no Brasil . . . . .	20
2.1.2 Processo de produção do biodiesel . . . . .	21
2.1.3 Qualidade do biodiesel . . . . .	23
2.1.4 Propriedades físico químicas relacionadas a qualidade do biodiesel . . . . .	24
2.2 Redes Neurais Artificiais . . . . .	27
2.2.1 Estruturas de RNAs . . . . .	29
2.2.2 Aprendizado em RNAs . . . . .	32
2.2.3 Aprendizagem por correção de erro e o algoritmo backpropagation . . . . .	34
2.3 Trabalhos relacionados . . . . .	38

<b>3</b>	<b>Abordagem de Desenvolvimento de RNAs Proposta para Predição de Propriedades de Biodiesel</b>	<b>42</b>
3.1	Processo de construção de RNAs . . . . .	42
3.1.1	Pré-treinamento e ferramenta utilizada . . . . .	43
3.1.2	Dados experimentais . . . . .	44
3.1.3	Pré-processamento dos dados . . . . .	45
3.2	Aspectos de especificação topológica das redes . . . . .	46
3.2.1	Definição de topologias candidatas . . . . .	46
3.2.2	Seleção de melhores topologias candidatas . . . . .	47
<b>4</b>	<b>Resultados e Discussões</b>	<b>50</b>
4.1	Considerações sobre o critério de parada antecipada . . . . .	50
4.2	Resultados obtidos para índice de iodo . . . . .	52
4.2.1	Resultados obtidos com uma camada . . . . .	52
4.2.2	Resultados obtidos com duas camadas . . . . .	53
4.3	Resultados obtidos para viscosidade cinemática . . . . .	57
4.3.1	Resultados obtidos com uma camada . . . . .	58
4.3.2	Resultados obtidos com duas camadas . . . . .	62
<b>5</b>	<b>Considerações Finais</b>	<b>71</b>
5.1	Conclusões do trabalho . . . . .	71
5.2	Limitações . . . . .	72
5.3	Principais contribuições . . . . .	72
5.4	Trabalhos futuros . . . . .	73
	<b>Referências Bibliográficas</b>	<b>75</b>

# 1 Introdução

Neste capítulo, serão apresentados o contexto envolvido do objeto de pesquisa, expondo a problemática a ser tratada, bem como, a solução proposta e os objetivos. A metodologia que foi empregada para a execução do trabalho de pesquisa também será explicitada, além da apresentação dos demais capítulos e organização desta dissertação.

## 1.1 Contexto

A busca mundial por alternativas limpas de energia, ou seja, aquelas geradas a partir de recursos renováveis e que não agravem os problemas de emissão de gases poluentes na atmosfera, fez com que vários países passassem a buscar alternativas e a realizar pesquisas direcionadas ao consumo de combustíveis de maneira sustentável. O Brasil, e o estado do Maranhão principalmente, se encontram em posição privilegiada por sua grande diversidade de recursos naturais para transformação em energia limpa e sustentável, no que se refere a sua potencialidade de produção de combustíveis derivados de fontes renováveis.

O biodiesel, por ser uma fonte energética oriunda da biomassa, é capaz de contribuir, com sustentabilidade, para suprir essa demanda energética mundial. Para assegurar uma combustão eficiente e qualidade nas emissões, bem como a segurança no transporte e manuseio do biodiesel, a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) institui, por meio da Resolução nº 14 de 2012 (atualizada recentemente para nº 45 em 2014) parâmetros de qualidade e especificações para esse biocombustível, e para isso diversos métodos oficiais ou alternativos podem ser utilizados. Na literatura, é possível identificar uma crescente utilização de métodos lineares e não-lineares no reconhecimento e classificação de padrões aplicados ao monitoramento da qualidade de biodiesel. Nesse contexto, pode-se citar trabalhos como em [1], [25], [31] que utilizam as Redes Neurais Artificiais (RNAs) como ferramentas não-linear na predição de propriedades de combustíveis.

O biodiesel apresenta diversas propriedades que podem estar relacionadas aos contaminantes da matéria-prima, ao processo produtivo, às condições de armazenamento e, principalmente, às estruturas moleculares que o compõem. As propriedades inerentes à composição, como a densidade, a viscosidade e o índice de iodo, estão diretamente relacionadas ao percentual de cada éster alquílico que compõe a matéria-prima (óleo vegetal ou gordura animal) e, conseqüentemente, o biodiesel, e apresentam um padrão de comportamento que viabiliza o seu estudo por meio de RNAs.

O presente trabalho propõe a avaliação da predição entre algoritmos de treinamento supervisionado de RNAs aplicadas ao monitoramento da qualidade de biodiesel, através da relação entre as composições de ésteres, que serão utilizadas para alimentar as redes, e as propriedades do biocombustível. Para se alcançar esse objetivo, é necessário avaliar os algoritmos de treinamento supervisionado de RNAs existentes na literatura, aplicar ao domínio da problemática comparando e validando quais apresentam os melhores resultados para predição dessas propriedades.

## 1.2 Problemática

A utilização do biodiesel como obrigatoriedade no Brasil entrou em vigor desde 1º de janeiro de 2010 com a Resolução nº 6 de 2009 do Conselho Nacional de Política Energética (CNPE), hoje o óleo diesel comercializado em todo o Brasil contém 7% de biodiesel. O Brasil tem um grande potencial para produção de biodiesel e está entre os maiores produtores e consumidores do mundo segundo a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP).

Para garantir a qualidade do biodiesel é necessário estabelecer padrões de qualidade, objetivando fixar teores dos contaminantes que não venham prejudicar a qualidade de emissões de queima, bem como desempenho, a integridade do motor e a segurança do transporte e manuseio [30]. Esses padrões para regulação e a fiscalização do biodiesel produzido no Brasil também são de responsabilidade da ANP e, em vista disso, por meio de resolução nº 14 de 2012, foram estabelecidos vários parâmetros de qualidade para esse biocombustível, e no ano de 2014 foi atualizado para resolução nº 45.

Vários métodos têm sido utilizados para verificação da qualidade de biodiesel, muito deles podem ser bastante demorados e onerosos para a indústria [30]. As redes neurais se apresentam como uma alternativa não-linear viável para monitoramento dos parâmetros de qualidade.

### 1.3 Solução Proposta

No treinamento de redes neurais, diversos algoritmos podem ser utilizados para a otimização de arquiteturas que possibilitem a convergência dos resultados de forma mais rápida e eficiente, como os métodos de Broyden-Fletcher-Goldfarb-Shanno, de gradiente conjugado escalonado, de Levenberg-Marquardt, entre outros. Em vista disso, a aplicação desses algoritmos no domínio do biodiesel torna-se uma possibilidade viável para monitoramento de propriedades qualidade.

Por meio da predição das propriedades de biodiesel a partir da composição de ésteres da matéria-prima, será possível avaliar a viabilidade da utilização de tal matéria-prima para a síntese de um biodiesel de qualidade. Além disso, alguns métodos de análise das propriedades de biodiesel bastantes demorado e/ou complexos poderão ser dispensados.

Desse modo, o presente trabalho visa verificar qual será a tendência da qualidade do biodiesel ou até mesmo dispensar o processo de síntese do biocombustível no caso de uma matriz inapropriada para essa finalidade, o que possivelmente renderá custos e tempos para as indústrias de biocombustíveis.

### 1.4 Objetivos

O objetivo geral deste trabalho é avaliar a predição de propriedades de qualidade de biodiesel utilizando algoritmos de treinamento de aprendizado supervisionado de RNAs.

### 1.4.1 Objetivos Específicos

- Identificar o estado da arte acerca do tema de domínio das áreas de biodiesel e suas propriedades, além das áreas de inteligência artificial, redes neurais artificiais aplicadas a biocombustíveis, aprendizagem de máquina, algoritmos de treinamento;
- Simular os algoritmos de treinamento aplicados a predição das propriedades de biodiesel;
- Avaliar quais algoritmos de treinamento apresentam os melhores resultados aplicadas ao monitoramento da qualidade do biodiesel;

## 1.5 Metodologia de pesquisa

Para alcançar os objetivos do trabalho, foram delimitadas e realizadas as seguintes atividades metodológicas:

1. Coleta de material bibliográfico para o estudo de caso e fundamentação da proposta da pesquisa.
2. Estudo acerca do domínio das áreas de biodiesel e suas propriedades, além das áreas de inteligência artificial, redes neurais artificiais aplicadas a biocombustíveis, aprendizagem de máquina, algoritmos de treinamento.
3. Estudo das ferramentas utilizadas para desenvolvimento e treinamento das RNAs da pesquisa.
4. Seleção dos algoritmos de treinamento supervisionado.
5. Implementação e testes dos algoritmos de treinamento mais adequados na predição das propriedades de biodiesel.
6. Comparação e otimização das arquiteturas de rede neurais em vista dos resultados obtidos.
7. Validação dos testes e análise dos resultados obtidos da aplicação de redes neurais.

## 1.6 Apresentação da Dissertação

A dissertação descrita sobre trabalho de pesquisa está organizada em 5 (sete) capítulos, como segue:

O primeiro Capítulo trata do contexto tecnológico e ambiental no qual todo o trabalho de pesquisa se insere. O capítulo também expõe a problemática que motivou a pesquisa, buscando o desenvolvimento de uma solução viável para predição de propriedades relativas a qualidade do biodiesel, no caso o índice de iodo e viscosidade. Neste primeiro capítulo, o objetivo geral, os objetivos específicos e a metodologia de pesquisa são apresentados.

O segundo Capítulo, a fundamentação teórica que norteia este trabalho é apresentada. Os principais conceitos necessários para o entendimento da solução proposta neste trabalho são expostos, tais como: O biodiesel, o desenvolvimento desse tipo de combustível no Brasil, o processo de produção utilizado para sua fabricação, a importância do controle de qualidade do biodiesel, os parâmetros escolhidos para determinar a qualidade de tal biodiesel. Além de demonstrar os conceitos de Redes Neurais Artificiais (RNAs) e suas principais características para resolução da problemática e o estado da arte relacionado ao contexto da pesquisa.

O terceiro Capítulo apresenta a abordagem metodológica de simulação das redes propostas para a predição de índice de iodo e viscosidade, o critério de escolha das topologias candidatas que melhor se aplicaram para a resolução do problema.

O quarto Capítulo apresenta as discussões e os resultados obtidos pelas configurações e simulações testadas, tanto para índice de iodo e para viscosidade na predição das RNAs.

Por fim, o quinto Capítulo traz as conclusões da pesquisa, suas contribuições, as limitações, dificuldades enfrentadas e os trabalhos futuros para aprimoramento da pesquisa iniciada.

## 2 Fundamentação Teórica

Este capítulo apresenta a base teórica necessária para o entendimento das tecnologias utilizadas no desenvolvimento da pesquisa.

### 2.1 Biodiesel

A poluição trazida devido ao uso de combustíveis fósseis e a ameaça iminente de escassez deste tipo de recurso, levou uma preocupação aos países de buscar fontes rentáveis de energia tanto social como economicamente. Essa crescente demanda energética mundial por fontes de energia limpa e sustentáveis fez do biodiesel um biocombustível chave para atender estas características juntamente com outras fontes renováveis.

O biodiesel é conceituado como uma mistura de aquil ésteres de ácidos graxos de cadeia longa, que podem ser derivadas de diferentes matérias primas, como óleos vegetais e gorduras animais, por meio da reação de transesterificação com algum álcool, que é o meio de produção mais utilizado [11,28,31]. Esse combustível substitui total ou parcialmente o diesel de petróleo em motores ciclo diesel de caminhões, tratores, camionetas, automóveis e também para geração de energia e calor. E pode ser usado puro ou misturado ao diesel em diversas proporções [8].

O biodiesel como substituto do óleo diesel comum apresenta diversas vantagens, dentre elas pode-se citar que é uma energia limpa, pois reduz a emissão de dióxido de carbono (CO<sub>2</sub>) em mais de 80%, produz 100% menos dióxido de enxofre (SO<sub>2</sub>) do que combustíveis derivados de petróleo e reduz o número de partículas sólidas poluentes em até 75%, mesmo quando associados aos outros combustíveis.

O biodiesel pode ser obtido de uma grande variedade de matérias-primas sendo que esses recursos são renováveis o que torna a degradação do biodiesel na natureza 4 vezes mais rápida que os derivados de petróleo; e também não requer mudanças muito significativas dos veículos automotores para poder ser utilizado nos atuais motores diesel. Porém alguns problemas ainda são associados ao uso do biodiesel,

como seu preço mais elevado que em muitos países é compensado pelos incentivos legislativos e regulamentares ou subsídios fiscais em forma de impostos especiais de consumo [28].

### 2.1.1 Biodiesel no Brasil

A trajetória do biodiesel no Brasil teve início em 2 de julho de 2003 por intermédio de decreto presidencial pelo qual foi instituído um grupo de trabalho interministerial com o objetivo de estudar a viabilidade do uso desse combustível como fonte alternativa de energia. O estudo mostrou então que o Brasil poderia investir neste tipo de combustível de cunho estratégico e economicamente viável, pois, as perceptivas de sua inserção na matriz energética brasileira direcionava a diversos benefícios sociais, ambientais e econômicos [17].

A partir desse estudo, O Programa Nacional de Produção e uso do Biodiesel - PNPB, nasceu em 2004, com o compromisso de viabilizar a produção e o uso do biodiesel no país, com foco na competitividade, na qualidade do biocombustível produzido, na garantia de segurança de seu suprimento, na diversificação de matérias primas, no fortalecimento das potencialidades regionais para produção, e, prioritariamente, na inclusão social de agricultores familiares [10].

O PNPB aferiu então uma gama de benefícios que a extensão continental do país proporciona e a diversidade geográfica que garante a disponibilidade de inúmeras fontes de matérias-primas para seleção de maior produção de biocombustíveis para alcançar os objetivos do projeto. Ofereceu um modelo de inclusão social econômica com a produção de matérias-primas por agricultura familiar, estimulando a inclusão social em todas as regiões do país. E definiu um modelo de produção de qualidade de biodiesel desde a obtenção da matéria-prima, produção e aproveitamento de biodiesel e seus subprodutos, logísticas para garantir preços competitivos e suprimentos.

Desde a implantação do PNPB, o programa é reconhecido dentro e fora do País como um sucesso econômico. Porém ainda existem muitos desafios a serem superados, como aproximar seu custo ao valor do óleo diesel fóssil, que ele deverá substituir, deixando progressivamente a suas externalidades como um legado sem uma conta a ser quitada pela economia do País. Assim, a reformulação dos modelos de produção e as matérias-primas deverão ser adaptadas para vencer este desafio [17].

Diante desse panorama, como forma de diversificar a matriz energética brasileira e reduzir a dependência de combustíveis fósseis o governo brasileiro, em 13 de janeiro de 2005, a Lei nº 11097 introduziu o biodiesel na matriz energética do país, sendo estabelecido 5% em volume, o percentual mínimo obrigatório de adição de biodiesel ao óleo diesel comercializado ao consumidor final, em qualquer parte do território nacional [9].

De acordo com o artigo 7 da mesma lei, ficou também instituída a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP o órgão responsável pela regulação, a contratação e a fiscalização da indústria do petróleo, gás natural, seus derivados e biocombustíveis. Assim, a ANP por meio da resolução nº 14 de 11 de maio de 2012 estabeleceu as obrigações e especificações quanto ao controle de qualidade a serem atendidas pelos diversos agentes econômicos que comercializam o produto em todo o território nacional, sendo atualizado pela resolução nº 45 em 26 de agosto de 2014.

Também, em agosto de 2014, foi aprovada pelo governo a Medida Provisória 647/14 que estabeleceu um aumento ao percentual obrigatório de mistura do biodiesel ao óleo diesel, que subiu para 6% em 1º julho e em seguida para 7% a partir de 1º de novembro desse ano. Até a edição da MP 647, o percentual era de 5% [48].

### **2.1.2 Processo de produção do biodiesel**

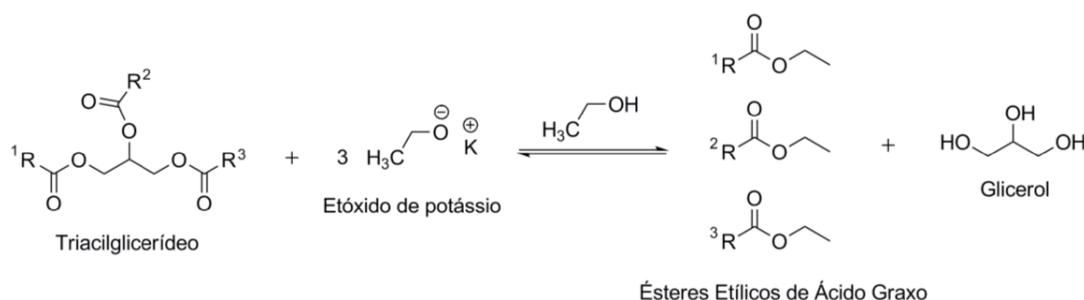
Para que óleos vegetais ou gorduras animais pudessem ser usados como combustíveis, estes devem ser convertidos em alquil ésteres de ácidos graxos, também chamado de biodiesel. O motivo pelo qual essa transformação deva acontecer é pela alta viscosidade dessas matérias graxas que não transesterificadas podem conduzir a sérios problemas operacionais nos motores diesel, tais como a ocorrência de depósitos em várias partes do motor.

Quatro métodos têm sido investigados para reduzir a alta viscosidade de óleos vegetais: uso de misturas binárias com petrodiesel, pirólise, microemulsificação (ou mistura co-solvente) e transesterificação. A transesterificação é de longe o método mais comum e o único que leva a produtos denominados de biodiesel [28].

Para se obter biodiesel, nesta reação, óleos vegetais e gorduras animais reagem na presença de um catalisador (geralmente uma base) com um álcool (geralmente metanol) para produzir alquil ésteres correspondentes (ou para metanol, metil ésteres) da mistura de ácidos graxos encontrados no óleo vegetal e gordura animal de origem, além de produzir glicerol como sub-produto da reação [3].

Em alguns países da Europa e nos Estados Unidos, por questões econômicas e por razões relacionadas com o processo, o metanol é o mais empregado na reação de transesterificação. O metanol apresenta vantagens sobre o etanol, pelo fato de apresentar menor teor de água, maior velocidade de reação e excelente rendimento na formação de ésteres graxos [16].

No Brasil, onde o etanol é mais barato, ésteres etílicos são utilizados como combustível [28]. O uso do etanol é extremamente importante, uma vez que tem-se a maior produção mundial. A utilização de etanol é atrativa do ponto de vista ambiental por ser produzido a partir de fonte renovável como a cana de açúcar e, ao contrário do metanol, apresenta baixa toxicidade e torna o biodiesel do ponto de vista ecológico mais adequado. Como desvantagem promove uma maior dispersão da glicerina no biodiesel (separação de fases é menos definida) e a reação de transesterificação é mais lenta, podendo formar emulsões durante o progresso da reação de transesterificação [16]. A Figura 2.1 mostra a reação de transesterificação, que é considerado um dos métodos mais comuns para obtenção de biodiesel. Aqui o triacilglicerídeo reage com um álcool na presença de catalisador ácido ou básico, resultando em moléculas menores de ésteres de ácidos graxos e glicerol.



**Figura 2.1:** Esquema da reação de transesterificação de um triacilglicerídeo catalisada por base.

Fonte: [16]

### 2.1.3 Qualidade do biodiesel

O atendimento a padrões de qualidade é feito objetivando fixar teores limites de contaminantes que não venham a prejudicar a qualidade das emissões da queima, bem como o desempenho, a integridade do motor e a segurança no transporte e manuseio. Quando as especificações são atendidas, o biodiesel pode ser utilizado na maioria dos motores modernos, sem exigir qualquer modificação ou oferecer qualquer comprometimento da durabilidade e da confiabilidade do motor. E mesmo quando utilizado em mistura ao diesel de petróleo, o biodiesel deve atender às especificações, independentes dos teores misturados ao diesel [28,30].

Várias normas internacionais estabelecem um padrão apropriado, como a European Standard (EN) 14214 na Europa, a American Society for Testing and Materials (ASTM) D6751 nos Estados Unidos [31]. No Brasil a resolução nº 14 da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) controla as obrigatoriedades de qualidade do biodiesel, mediante o emprego das normas da Associação Brasileira de Normas Técnicas (ABNT), e das normas internacionais ASTM, da International Organization for Standardization (ISO) e do Comité Européen de Normalisation (CEN) [11].

O controle de qualidade do biodiesel é importante, pois, a sua qualidade pode ser influenciada por vários fatores, incluindo a qualidade da matéria-prima, a composição em ácidos graxos do óleo vegetal ou gordura de origem animal, o processo de produção, o emprego de outros materiais e parâmetros posteriores a produção [28]. No que tange as estruturas moleculares dos ésteres podem variar tanto no tamanho da cadeia carbônica, quanto na quantidade e posição de insaturações. Contaminantes procedentes da matéria-prima, a exemplo do fósforo, enxofre, cálcio e magnésio, podem também ser encontrados no biodiesel [30]. Em relação a fatores posteriores a produção, água e materiais em suspensão, como ferrugem ou partículas de sujeira, podem ser encontrados no biodiesel, ou ainda podem se originar da oxidação do combustível mediante a formação de compostos insolúveis [28].

Diante do exposto garantir um combustível de alta qualidade é um pré-requisito essencial para que o biodiesel tenha aceitação do mercado, já que a manutenção de sua oferta por meio do atendimento às exigências de especificações faz do

biodiesel um importante aliado que contribui com tripé da sustentabilidade: ambiental, social e econômica.

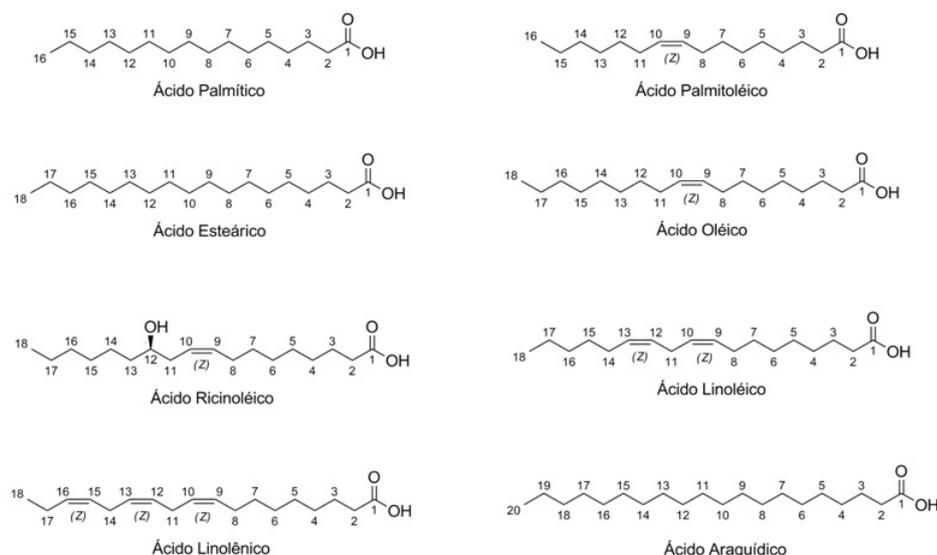
#### **2.1.4 Propriedades físico químicas relacionadas a qualidade do biodiesel**

O biodiesel apresenta propriedades físico-químicas que podem variar dependendo do seu óleo de origem, processo de produção ou manuseio durante estocagem. Além dessas, algumas propriedades do biodiesel estão relacionadas com as estruturas moleculares dos seus alquil ésteres constituintes [22, 28, 30].

Ao considerar a matéria-prima, quando são oleaginosas, alguns itens devem ser levados em consideração, tais como: composição química do óleo extraído, percentual de óleo na semente, elevado rendimento energético por unidade de área, produção abundante de óleos a custos reduzidos, além do tipo e também sua disponibilidade em determinada região. A composição química do óleo ou gordura é essencial, visto que influencia nas propriedades físicas do óleo, como viscosidade, o índice de cetano, entre outros. Podendo também servir de base para uma análise preliminar sobre a viabilidade ou não, de dada matéria-prima para a produção de um biocombustível que atenda completamente os parâmetros de qualidade da norma vigente [16].

Aproximadamente 90% dos óleos vegetais produzidos são formados por triacilglicerídeos de ácidos graxos de cadeia longa, que varia de 16 a 20 átomos de carbono, tendo como principais constituintes os ácidos: esteárico (C18:0), oléico (C18:1), linoléico (C18:2) e linolênico (C18:3) [16]. A Figura 2.2 apresenta as estruturas mais comuns de ácidos graxos que estão presentes em óleos vegetais utilizados para produção de biodiesel.

As propriedades dos vários ésteres de ácidos graxos são determinadas pelas características estruturais do ácido graxo e pelas porções de álcool que compreendem um éster graxo. Uma vez que a reação de transesterificação de um óleo ou gordura leva a um biodiesel, o seu perfil de ácido graxo corresponde ao perfil do ácido graxo do óleo ou gordura-mãe. Essas características estruturais que influenciam as propriedades físicas e de combustível de uma molécula de éster graxo são o comprimento de cadeia, grau de insaturação, e ramificação da cadeia [28].



**Figura 2.2:** Estruturas de ácidos graxos mais comuns presentes em óleos vegetais. Fonte: [16]

Assim, importantes propriedades de combustível de biodiesel, que são influenciados pelo perfil de ácidos graxos e, por sua vez, pelas características estruturais dos vários ésteres de ácidos graxos são utilizados para o controle de qualidade e desempenho do combustível. Podemos citar densidade, viscosidade cinemática, número de cetano, índice de iodo, ponto de névoa, ponto de entupimento de filtro a frio e ponto de fluidez. Para o contexto do trabalho iremos discorrer um pouco mais sobre a propriedade índice de iodo e viscosidade que foram as propriedades preditivas do trabalho.

### Índice de Iodo

O índice de iodo (II) é uma medida de insaturação total dentro de uma mistura de ácidos graxos, baseada na adição formal de iodo nas duplas ligações e determinada em g de iodo/100g de amostra. O II do biodiesel é praticamente idêntico ao valor do óleo utilizado na sua produção [28, 39]. Este valor influencia em várias propriedades importantes do biodiesel, como ponto de entupimento de filtro a frio e a estabilidade oxidativa. Óleos e biodiesel com elevado valor de iodo tendem, a ser menos estáveis e mais suscetíveis à oxidação, devido ao elevado grau de insaturação [1,3]. O uso do II é justificado pois este parâmetro indica a propensão de óleo ou gordura a oxidação, além de indicar a propensão do óleo ou gordura a polimerização e formação de depósitos no motor [28].

A determinação dessa propriedade é feita através de métodos oficiais desenvolvidos pelo American Oil Chemists' Society (AOCS) e a Association of Analytical Communities. O método analítico EN ISO 14111 adotado pelas principais normas internacionais, baseia-se no tratamento da amostra com halogênios em excesso, que se adicionarão às duplas ligações. Os halogênios não reagidos são então titulados como tiosulfato de sódio e o resultado expresso como gramas de iodo que reagem com as insaturações em 100 g de amostra. Também é possível calcular o número de iodo e a massa molar média de biodieseis etílicos utilizando a espectroscopia de ressonância nuclear magnética.

Contudo, esses métodos tendem a ser trabalhosos, demorados e envolvem vários passos usando diversos solventes, e produtos químicos tóxicos. Muitos métodos apresentam elevado custo de aquisição de manutenção de equipamentos [1, 30, 34]. Assim, vários métodos livres e não invasivos estão sendo desenvolvidos para determinar ou prever o número de iodo em [1, 3, 23, 34].

### **Viscosidade Cinemática**

A viscosidade é uma das propriedades mais importantes que afetam a utilização do biodiesel como combustível [2]. É uma propriedade que possui relação direta com o processo de injeção de combustível para motores. A viscosidade do biodiesel aumenta com o comprimento da cadeia carbônica e com o grau de saturação [28].

A presença de sabões residuais, triacilglicerídeos e os produtos da degradação oxidativa no biodiesel, aumentam a viscosidade ocasionando a deposição de resíduos nas partes internas do motor [30]. Estes contaminantes podem, portanto, ser monitorados indiretamente através da determinação da viscosidade cinemática a 40 °C. A norma EN 14214 (EN ISO 3104) estabelece um intervalo aceitável de 3,5 a 5,0 mm<sup>2</sup>/s, ASTM D6751 (D 445) permite um intervalo de 1,9 a 6,0 mm<sup>2</sup>/s e a ANP com limite aceitável de 3,0 a 6,0 mm<sup>2</sup>/s. A ANP além dos métodos analíticos citados recomenda também o método ABNT NBR 10441. [16]

A previsão da viscosidade cinemática tem recebido atenção considerável na literatura. Tanto a viscosidade do óleo e seu éster alquílico derivado pode ser utilizado para monitorar a qualidade do biodiesel [2]. Uma vez que o processo para a obtenção de biodiesel, que a transesterificação dos ésteres de ácidos graxos, é utilizada com o

objetivo de reduzir a elevada viscosidade dos óleos vegetais, e assim, permitir a sua utilização em combustíveis para motores diesel sem problemas operacionais.

## 2.2 Redes Neurais Artificiais

Redes neurais artificiais (RNAs) são uma tecnologia que têm sido utilizada para diversos fins, baseada no funcionamento dos neurônios biológicos, possui grande capacidade de aprendizado e generalização diante de exemplos expostos, mesmo com a interferência de ruídos [21]. Conhecida como máquinas de processamento paralelo e por ser uma técnica de estatística não-linear, pode resolver uma grande variedade de problemas e pelo seu caráter multidisciplinar está nas mais diversas aplicações nas quais podemos citar: avaliação de imagens captadas por satélite, classificação de padrões de escrita e fala, reconhecimento de faces em visão computacional, controle de trens de grande velocidade, previsão de ações no mercado financeiro, identificação de anomalias em imagens médicas [4,21,26].

Inseridas na área de pesquisa conhecida como Inteligência Artificial [26], mais precisamente na abordagem conexionista, as RNAs podem ser generalizadas por modelos matemáticos que assumem que o processamento das informações ocorre em vários elementos chamados neurônios; os sinais são propagados de um elemento a outro através das conexões; cada conexão possui um peso associado, que, em uma rede neural típica, pondera o sinal transmitido e cada neurônio (ou unidade) aplica uma função de ativação, que em geral é não-linear, à sua entrada de rede para determinar sua saída [47].

A analogia entre o comportamento do neurônio biológico, que tem seu modelo simplificado na Figura 2.3, e o neurônio artificial foi o que possibilitou e deu base para as pesquisas em redes neurais. Fazendo a correlação com o modelo do neurônio mais simples de McCulloch & Pitts (1943), os sinais de entrada representados pelo conjunto  $\{x_1, x_2, \dots, x_n\}$ , equivalem aos impulsos elétricos externos captados pelos dendritos.

As ponderações exercidas pelas junções sinápticas do modelo biológico são representadas no neurônio artificial pelo conjunto de pesos sinápticos  $\{w_1, w_2, \dots, w_n\}$ . Obtendo-se a soma ponderada dos sinais de entrada  $\{x_n\}$  e seus respectivos pesos

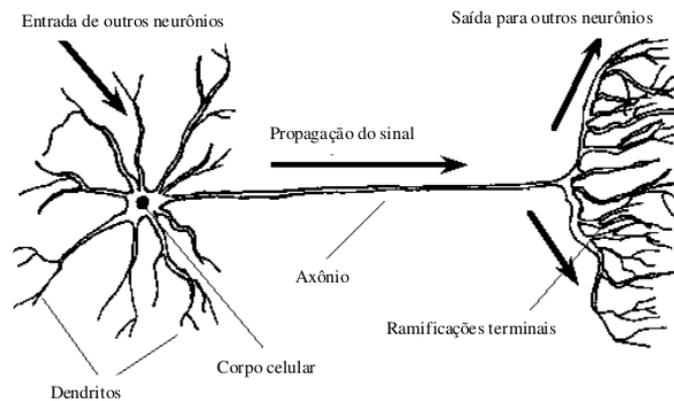


Figura 2.3: Modelo da célula neural biológica. Fonte: [47]

sinápticos  $\{w_n\}$  obtemos o potencial de ativação, denotado por  $u$ , e representado no neurônio biológico pelo corpo celular que gera o sinal de saída propagado pelo axônio. Esse sinal de saída é representado por  $y$  no neurônio artificial.

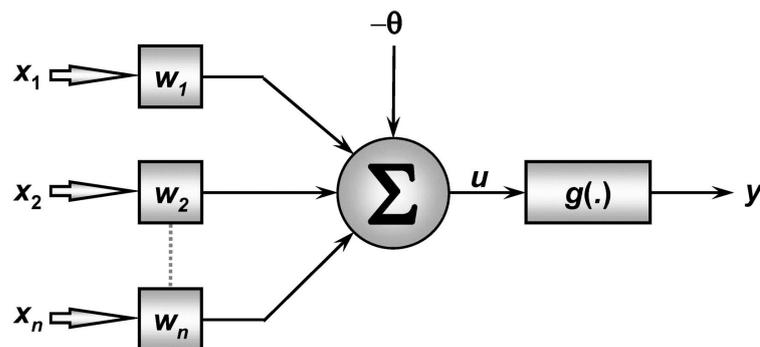


Figura 2.4: Modelo da célula neural artificial. Fonte: [15]

Considerando a Figura 2.4, identifica-se os principais elementos que constituem o neurônio artificial, que de acordo com [15] são 7:

- Sinais de entrada  $\{x_1, x_2, \dots, x_n\}$

São sinais advindos do meio externo e que representam os valores assumidos pelas variáveis de uma aplicação específica.

- Sinais de saída  $\{w_1, w_2, \dots, w_n\}$

São os valores que servem para ponderar cada uma das variáveis de entrada da rede, permitindo-se quantificar a sua relevância em relação à funcionalidade do respectivo neurônio.

- Combinador linear  $\{\Sigma\}$

Sua função é agregar todos os sinais de entrada e ponderá-los com os respectivos pesos sinápticos a fim de produzir um valor de potencial de ativação.

- Limiar de ativação  $\{\theta\}$

É uma variável que especifica qual será o patamar apropriado para que o resultado produzido pelo combinador linear possa gerar um valor de disparo em direção a saída do neurônio.

- Potencial de ativação  $\{u\}$

É o resultado produzido pela diferença entre o combinador linear e o limiar de ativação. O resultado de  $u$  é representado por:

$$u = \sum_{i=0}^n w_i \cdot x_i - \theta \quad (2.1)$$

- Função de ativação  $\{g\}$

A função limita a saída do neurônio dentro de um intervalo de valores razoáveis a serem assumidos pela sua própria imagem funcional. Dependendo do domínio empregado as redes devem apresentar uma não – linearidade em sua saída, que seja suave, portanto, diferenciável. Então, quando propriedades dinâmicas estão envolvidas na definição do estado de ativação, equações diferenciais ou as diferenças são empregadas [47]. As principais funções utilizadas são demonstradas na Figura 2.5, onde  $\beta$  é uma constante real associada ao nível de inclinação da função logística frente ao seu ponto de inflexão.

- Sinal de saída  $\{y\}$

É o valor final produzido pelo neurônio em relação a um determinado conjunto de sinais de entrada, podendo ser também utilizado por outros neurônios que estão sequencialmente interligados.

$$y = g(u) \quad (2.2)$$

### 2.2.1 Estruturas de RNAs

A maneira pela qual os neurônios de uma rede neural são estruturados está intimamente ligada com o algoritmo de aprendizagem usado para treinar a rede.

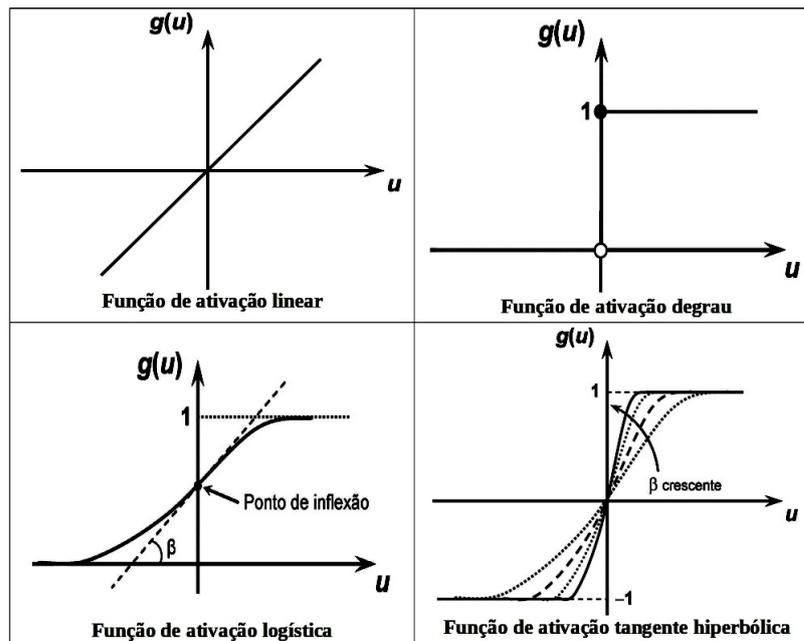


Figura 2.5: Principais funções de ativação. Fonte: [15]

Pode-se, portanto, falar de algoritmos de aprendizagem utilizados no projeto de redes neurais como sendo estruturados [21]. Assim, o desempenho de uma rede neural depende não só da função de ativação usada pelos neurônios, mas também de como esses neurônios são arranjados, isto é, estruturados uns com os outros, e dos algoritmos de aprendizado utilizados no processo de treinamento da rede.

Os autores em [26] consideram apenas dois tipos de arquiteturas de RNA: redes *feedforward* e redes recorrentes. Já os autores [15] e [21] dividem a arquitetura de uma RNA em três tipos: redes *feedforward* de camada simples, redes *feedforward* de camada múltiplas e redes recorrentes. Aqui consideraremos estes três tipos de arquitetura.

### Redes *feedforward* de camada simples

Este tipo de estrutura é formada por apenas uma camada de entrada e uma única camada de neurônios, que é a mesma camada de saída. Esta rede é estritamente do tipo alimentada adiante ou acíclica, e é chamada de rede de camada única, sendo que a designação "camada única" se refere à camada de saída de nós computacionais (neurônios). Não contamos a camada de entrada de nós, porque lá não é realizada qualquer computação [21].

Exemplos de redes desse tipo de arquitetura são as redes Perceptron e a Adaline, que pode ser observado na Figura 2.6.

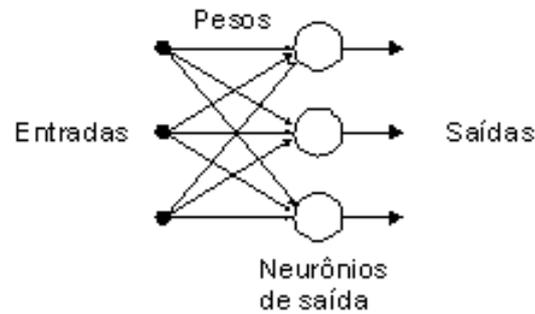


Figura 2.6: Rede *feedforward* de camada simples. Fonte: [49]

### Redes *feedforward* de camada múltipla

As redes *feedforward* de camadas múltiplas são constituídas pela presença de uma ou mais camadas escondidas de neurônios. Com os neurônios ocultos é possível intervir entre a camada de entrada e a saída da rede de uma maneira útil tornando a rede capaz de extrair estatísticas de ordem elevada. Como exemplos desse tipo de rede temos as perceptron múltiplas camadas (Multilayer Perceptron – MLP) e as redes de funções de base radial (radial basis function – RBF).

As redes do tipo MLP têm sido aplicadas com sucesso para resolver diversos tipos de problemas de cunho geral, tais como aproximação de funções, classificação, categorização e predição. Essa flexibilidade, fez com que este tipo de rede fosse aplicado em diversas áreas. Seu treinamento é realizado de forma supervisionada por meio do algoritmo de retropropagação de erro, o qual é baseado na regra de aprendizagem por correção de erro [6]. A Figura 2.7 demonstra um exemplo dessa rede.

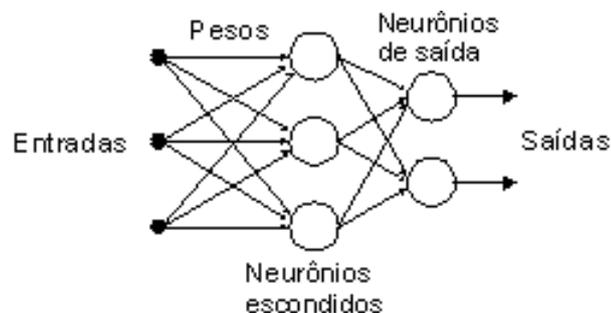


Figura 2.7: Rede *feedforward* de camada múltipla. Fonte: [49]

### Redes recorrentes

São redes em que as saídas dos neurônios são realimentadas como sinais de entrada para outros neurônios. Estas redes possuem uma característica de realimentação que as qualifica para o processo dinâmico de informações [15]. Um exemplo para esse tipo de rede são as redes Hopfield, que pode ser ilustrado na figura.

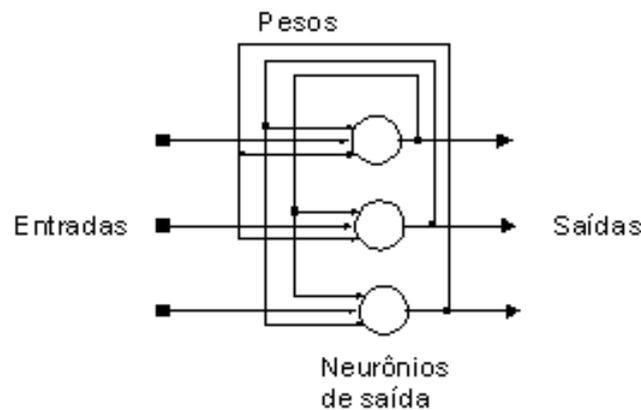


Figura 2.8: Rede recorrente. Fonte: [49]

### 2.2.2 Aprendizado em RNAs

Aprendizagem é uma das características mais relevantes no processo de treinamento de uma rede neural. A definição geral de aprendizagem é expressa da seguinte forma [21]:

“A aprendizagem é um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. O tipo de regra de aprendizagem é determinado pela maneira pela qual a modificação dos parâmetros ocorre.”

As regras de aprendizagem definem como ocorrerá o ajuste dos pesos entre os sucessivos ciclos de treinamento de uma rede neural [4]. Os tipos básicos de regra de aprendizagem são: aprendizagem por correção de erro; aprendizagem baseada em memória; aprendizagem hebbiana; aprendizagem competitiva e a aprendizagem de Boltzman.

Essas regras de aprendizagem para treinamento de redes neurais, segundo autores [7] e [21] agrupam-se em dois paradigmas principais: aprendizado supervisionado e aprendizado não – supervisionado. A referência [26] considera três grupos:

aprendizado supervisionado, aprendizado não-supervisionado e aprendizado híbrido. Aqui consideraremos a literatura de [21] para esta classificação. Apesar de ainda não se chegar a um consenso se a aprendizagem híbrida ou aprendizagem por reforço faz parte do paradigma supervisionado ou não-supervisionado.

### **Aprendizado supervisionado**

No aprendizado supervisionado é considerado em cada amostra sinais de entrada e sinais de saída desejada no processo. Também chamada de aprendizagem com professor, pois podemos considerar o professor como tendo conhecimento sobre o ambiente, sendo o conhecimento representado pelo conjunto de exemplos de entrada-saída.

Segundo [21] o professor é capaz de fornecer à rede neural uma resposta desejada, que representa a ação ótima a ser realizada pela rede neural. Os parâmetros da rede são ajustados sob a influência combinada do vetor de treinamento e do sinal de erro. O sinal de erro é definido como a diferença entre a resposta desejada e a resposta real da rede. O ajuste é então realizado com o objetivo da rede atingir a solução ótima. Quando a condição é alcançada, o professor deixa de atuar sobre a rede neural e lida por si mesma com o ambiente.

### **Aprendizado não – supervisionado**

O aprendizado não-supervisionado, chamado também de aprendizagem sem professor, somente os padrões de entrada estão disponíveis para a rede, ao contrário do aprendizado supervisionado, cujo conjunto de treinamento possui pares de entrada e saída [7].

Esse aprendizado subdivide-se em ainda duas seções [21]:

#### **a) Aprendizagem por reforço**

Em muitas literaturas este tipo de aprendizagem é considerado como um tipo de aprendizagem supervisionada ou como um terceiro paradigma de aprendizagem. A aprendizagem por reforço é realizada por meio de um crítico externo que procura maximizar o reforço das boas ações executadas pela rede.

O processo de treinamento da rede é realizado tipicamente por tentativa e erro, pois a única resposta disponível para determinada entrada é se esta é satisfatória ou não. Se for considerada satisfatória, incrementos nos pesos sinápticos e limiares são então gradualmente efetuados visando recompensar esta condição comportamental [15].

#### b) Aprendizagem não-supervisionada

Neste tipo de aprendizagem não existe qualquer professor ou crítico que ajude no treinamento da rede. Em vez disso, são dadas condições para a rede se auto organizar, de acordo com uma medida de representação e os parâmetros livres da rede são otimizados em relação a esta medida.

Durante o processo de aprendizado, os padrões de entrada são apresentados continuamente à rede, e a existência de regularidades nesses dados faz com que o aprendizado seja possível. Regularidade e redundância nas entradas são características essenciais para haver aprendizado não-supervisionado [7].

### 2.2.3 Aprendizagem por correção de erro e o algoritmo backpropagation

A aprendizagem pode ser conceituada como um processo pelo qual os parâmetros livres de uma rede neural são adaptados por meio de um processo de estimulação pelo ambiente no qual a rede está inserida. O tipo de aprendizagem é determinado pela maneira pela qual a modificação dos parâmetros ocorre [21].

A aprendizagem por correção de erro consiste basicamente em usar o sinal de erro para modificar os pesos de conexão para reduzir gradualmente este erro. Segundo o autor em [47] é considerado no processo de retro-propagação do erro duas fases: uma fase de propagação do sinal funcional (*feedforward*) e uma retro-propagação do erro (*feedback*). Na fase de propagação do sinal funcional, os vetores de dados são aplicados às unidades de entrada, e seu efeito se propaga pela rede, camada a camada. Finalmente, um conjunto de saídas é produzido como resposta da rede. Durante a fase positiva, os pesos das conexões são mantidos fixos. Na retro-propagação do erro, por outro lado, os pesos são ajustados com uma regra de correção do erro. Este sinal de

erro é propagado da saída para a entrada, camada a camada e os pesos são ajustados de forma que a distância entre a resposta da rede e a resposta desejada.

O desenvolvimento do algoritmo de aprendizado *backpropagation* para ajustar pesos em um perceptron de múltiplas camadas fez estas redes a mais popular entre os pesquisadores e usuários de redes neurais [26]. A equação que demonstra o método do *backpropagation* é:

$$\Delta w = -\eta \frac{\partial E_h}{\partial w_{hj}} \quad (2.3)$$

Para se obter um melhor desempenho do algoritmo, foi adicionado um termo chamado de *momentum* normalizado e que acrescenta o conhecimento da correção de erro feita no passo anterior na correção de peso atual. Desde modo, a equação final do algoritmo *backpropagation* ficou da seguinte forma apresentada [41]:

$$\Delta w_{hj} = -\eta \frac{\partial E(n)}{\partial w_{hj}(n)} + \alpha \Delta w_{hj}(n-l) \quad (2.4)$$

O algoritmo por retro-propagação de erro possibilitou sua aplicação a uma grande gama de problemas de classificação e mapeamento estatístico não-linear. Já que o processo de treinamento supervisionado de redes neurais artificiais perceptron *feedforward* é equivalente a um problema de otimização não-linear irrestrito, onde uma função de erro global é minimizada a partir do ajuste dos parâmetros da rede neural [47].

Entretanto, um dos problemas da implementação de uma rede MLP é que há casos que a velocidade de aprendizagem é um fator limitante para possibilitar a implementação prática deste tipo de ferramenta computacional no processo de solução de problemas que requerem otimalidade, robustez e rapidez na convergência do processo de ajuste de parâmetros [47].

Diversas variações têm sido propostas para o método *backpropagation* com o objetivo de tornar o processo de convergência mais rápido e eficiente. Entre os quais podemos citar os algoritmos:

- *Backpropagation* com *momentum* e taxa de aprendizagem:

Esta variação utiliza a inserção de um termo de *momentum* visando ponderar o quão as matrizes sinápticas foram alteradas entre duas iterações anteriores e sucessivas. Considerando os neurônios pertencentes à L-ésima camada de uma rede, tem-se:

$$W_{ji}^{(L)}(t+1) = W_{ji}^{(L)}(t) + \underbrace{\alpha \cdot W_{ji}^{(L)}(t) - W_{ji}^{(L)}(t-1)}_{\text{termo de momentum}} + \underbrace{\eta \cdot \delta_j^{(L)} \cdot Y_i^{(L-1)}}_{\text{taxa de aprendizagem}} \quad (2.5)$$

onde  $\alpha$  é a taxa de *momentum* e seu valor está compreendido entre zero e um. Pode - se abstrair que quando o valor da taxa de *momentum* for igual a zero, a expressão se torna equivalente àquela do *backpropagation* convencional. Por outro lado, para valores diferentes de zero, o termo de momentum passa a ser relevante, sendo que tal contribuição afetará positivamente o processo de convergência [15]. Assim, observa-se que a contribuição do termo de *momentum* quando se está perto da solução final é bem pequena, pois a partir desse instante, os ajustes de peso acabam ficando a cargo da taxa de aprendizagem, como ocorre no *backpropagation* convencional.

- *Backpropagation* BFGS Quase-Newton;

O método do BFGS é classificado como um método de Quase-Newton, onde a ideia aqui é fazer uma aproximação iterativa da inversa da matriz hessiana, utilizando informações de primeira ordem obtidas durante o processo iterativo de aprendizagem. A aproximação atual  $H_i$  é utilizada a cada iteração para definir a próxima direção descendente do método. Idealmente, as aproximações convergem para a inversa da matriz hessiana [47]. A expressão que permite determinar a aproximação da inversa da hessiana é apresentada na equação:

$$H_i + 1 = H_i + \frac{p_i p_i^T}{p_i^T q_i} \left[ 1 + \frac{q_i^T H_i q_i}{p_i^T q_i} \right] - \frac{H_i q_i p_i^T + p_i q_i^T H_i}{p_i^T q_i} \quad (2.6)$$

- Levenberg-Marquardt;

O algoritmo de Levenberg-Marquardt é uma aproximação do método de Newton, ele melhora o método de Gauss-Newton por meio da utilização de uma taxa de aprendizado variável [7]. Esse algoritmo apesar de demandar uma grande quantidade de memória e elevada complexidade computacional incorporado ao algoritmo *backpropagation* pode potencializar a eficiência do processo de treinamento. Os pesos da rede são ajustados de acordo com a equação 2.7, onde  $\nabla^2 E(w_{ji}(t))$  é uma matriz Hessiana e  $\nabla E(w_{ji}(t))$  é o gradiente.

$$\Delta w_{ji}(t) = -[\nabla^2 E(w_{ji}(t)) + \mu I]^{-1} \nabla E(w_{ji}(t)) \quad (2.7)$$

- *Resilient-Propagation;*

O método do Resilient-Propagation não considera as variações das magnitudes do gradiente da função de erro, levará somente em conta a variação do sinal, tornando a taxa de aprendizagem do método dinâmica [15]. Assim o algoritmo procura eliminar a influência negativa do valor da derivada parcial na definição do ajuste dos pesos. Já que essa influência ocorre porque, quando a saída de um neurônio for próxima de 0 (ou 1) e a saída desejada for 1 (ou 0), a derivada será próxima de 0, fazendo com que os pesos deste neurônio recebam um ajuste mínimo, próximo de 0 [7]. Em termos matemáticos, o ajuste dos pesos e avaliação de mudança de sinal é dada pela Equação 2.8 :

$$\Delta w_{ji}(t) = \begin{cases} -\Delta_{ji}(t), & \text{se } \frac{\partial E}{\partial w_{ji}}(t) > 0 \\ +\Delta_{ji}(t), & \text{se } \frac{\partial E}{\partial w_{ji}}(t) < 0 \\ 0, & \text{se } \frac{\partial E}{\partial w_{ji}}(t) = 0 \end{cases} \quad (2.8)$$

- Gradiente Conjugado Escalonado;

Os métodos do gradiente conjugado possuem sua estratégia baseada no modelo geral de otimização apresentado no algoritmo padrão. Já os métodos de segunda ordem até agora apresentados utilizam um procedimento de busca unidimensional para a determinação da taxa de aprendizagem. A busca unidimensional envolve um grande número de avaliações da função ou de sua derivada, tornando o processo bastante custoso do ponto de vista computacional [47].

O método do gradiente conjugado escalonado é uma variação que evita essa busca unidimensional por meio da utilização da abordagem do Levenberg-Marquardt ao fazer um escalonamento do passo de ajuste em  $\alpha$ . A idéia então é fazer a combinação da abordagem do gradiente conjugado e como Levenberg-Marquardt resultando na Equação 2.9 para aplicável as redes MLP, onde  $s_j$  é a informação de segunda ordem para o cálculo da matriz hessiana:

$$s_j = \frac{\nabla J(\theta_j + \sigma_j d_j) - \nabla J(\theta_j)}{\sigma_j} + \lambda_j d_j \quad (2.9)$$

- *One-Step-Secant*;

O método da secante provém do fato de que as derivadas são aproximadas por secantes avaliadas em dois pontos da função (neste caso a função é o gradiente). A vantagem desse método é que sua complexidade é de ordem  $O(P)$ , ou seja, é linear em relação ao número  $P$  de parâmetros, enquanto a complexidade de métodos como por exemplo, o BFGS, é de ordem  $O(P^2)$  [47]. A principal razão de redução da complexidade é que a direção de atualização  $d_i$  é calculada somente a partir de vetores determinados pelos gradientes, e não há mais a armazenagem da aproximação da inversa da hessiana:

$$d_{i+1} = -g_i + A_i s_i + B_i q_i \quad (2.10)$$

## 2.3 Trabalhos relacionados

Estudos sobre a influência da composição química do biodiesel nas suas propriedades de qualidade têm sido demonstrada por meio de pesquisas ressaltando que as propriedades dos triglicéridios e do combustível biodiesel são determinadas pelas quantidades de cada ácido graxo que se encontram presentes em suas moléculas [39].

Diferentes técnicas de previsão de propriedades de qualidade como índice de iodo, viscosidade cinemática, densidade, número de cetano, entre outros são foco de análise. Algumas abordagens usam modelos empíricos e estatísticos [2,14,18,19,38], além de técnicas que utilizam espectroscopia de infravermelho próximo (NIR) para determinar as propriedades do biodiesel [1,3]. As RNAs, por sua vez, têm sido aplicadas em várias áreas da engenharia e domínios. Em relação a pesquisas relacionadas a medida do controle de qualidade de combustíveis e biocombustíveis, essa ferramenta não-linear está sendo utilizada nos mais variados contextos.

Pode-se citar o trabalho de R. Piloto-Rodríguez et al. [36], onde as RNAs determinaram o número de cetano, um dos indicadores de qualidade do biodiesel ligado as propriedades dos ésteres. Essa propriedade é importante pois quanto maior o número de cetano mais fácil sua ignição. As RNAs aqui utilizadas foram de duas topologias para treinamento: (11:5:1) e (11:7:1), com 48 amostras de biodiesel e 15 para

teste. Segundo o autor, o modelo de predição a partir de redes neurais foi mais satisfatório comparado a regressão múltipla linear, porém o modelo não é recomendado para prever o número de cetano de ésteres de ácido graxo puro diferente dos 10 selecionados para o trabalho e a melhor performance obtida foi com a topologia (11:5:1), com índice de correlação de 0.95 para etapa de validação.

Jahirul et al. [25] utilizou redes neurais para predição da viscosidade cinemática a partir da influência da temperatura. A viscosidade é uma das propriedades mais críticas para a qualidade de um biocombustível, e é altamente dependente da temperatura para ter um comportamento adequado no momento da combustão. O trabalho utiliza 204 amostras com 23 entradas incluindo temperatura e composição química do biodiesel e a propriedade da viscosidade cinemática como saída. As RNAs possuem camada oculta variando de 23 a 46 neurônios, sendo as redes treinadas somente com o algoritmo Levenberg Marquardt. A melhor performance obtida foi com a topologia (23:28:1) com índice de determinação ( $R^2$ ) de 0.97, por outro lado a acurácia de predição do modelo pode ser melhorado através do aumento do intervalo de treinamentos e conjunto de dados.

Meng et al. [31] também conduziu seu trabalho com predição da viscosidade cinemática, neste estudo, uma RNA foi desenvolvida para prever a viscosidade cinemática do biodiesel à temperatura de 313 K, com os dados experimentais de 105 amostras de biodiesel recolhidas a partir da literatura. O método utiliza na RNA apenas as frações de massa de 19 ésteres metílicos de ácidos graxos, como entradas. Dois métodos previamente descritos baseados em equações empíricas são comparados com os resultados da RNA. Os resultados indicam que o método RNA proposto é capaz de prever a viscosidade cinemática a 313 K do biodiesel com o erro quadrado médio (MSE) de 0.0099. No geral, o método RNA demonstrou uma melhor precisão para a previsão da viscosidade cinemática biodiesel com o maior coeficiente de correlação de 0.9774.

O trabalho de Saldana et al. [42] utilizou métodos de aprendizagem de máquina que incluem Máquinas de Vetor de Suporte (SVM) e RNAs para predição de densidade e viscosidade de compostos de biocombustíveis a partir da influência de diferentes temperaturas, além de prever outras propriedades a partir da base de dados utilizada, a habilidade de previsão do modelo foram validadas com coeficiente de correlação de até 0.98 e 0.99 para RNAs e SVM respectivamente.

**Tabela 2.1:** Modelos desenvolvidos para predição de propriedades de qualidade de combustíveis e biocombustíveis.

Referência	Abordagem de predição
R. Piloto-Rodríguez et al. [36] Prediction of the cetane number of biodiesel using artificial neural networks and multiple linear regression	RNA e MRL
Y. Çay et al. [13] Prediction of engine performance and exhaust emissions for gasoline and methanol using artificial neural network	RNA
Jahirul. et al [25] An Artificial Neural Network (ANN) Model for Predicting Biodiesel Kinetic Viscosity as a Function of Temperature and Chemical Compositions	RNA
M. Rezaee et al. [40] A multivariate modeling for analysis of factors controlling the particle size and viscosity in palm kernel oil esters-based nanoemulsions	RNA
H. Sharon et al. [46] Biodiesel production and prediction of engine performance using SIMULINK model of trained neural network	RNA
Saldana et al. [42] Prediction of Density and Viscosity of Biofuel Compounds Using Machine Learning Methods	RNA e SVM
X. Meng et al. [31] Neural network prediction of biodiesel kinematic viscosity at 313 K	RNA
J. C. Chavarria-Hernandez e D. E. Pacheco-Catalán. [14] Predicting the kinematic viscosity of FAMES and biodiesel: Empirical models	Modelos empíricos e estatísticos
S.V.D. Freitas et al. [18] Measurement and prediction of high-pressure viscosities of biodiesel fuels	Modelos empíricos e estatísticos
C.A.W. Allen et al. [2] Predicting the viscosity of biodiesel fuels from their fatty acid ester composition	Modelos empíricos e estatísticos
M. Garcia et al. [19] Density of alkyl esters and its mixtures: A comparison and improvement of predictive models	Modelos empíricos e estatísticos
H. Sanli et al. [44] Predicting the higher heating values of waste frying oils as potential biodiesel feedstock	Modelos empíricos e estatísticos
L. F. Ramírez-Verduzco et al. [38] Predicting cetane number, kinematic viscosity, density and higher heating value of biodiesel from its fatty acid methyl ester composition	Modelos empíricos e estatísticos
P. Baptista. et al [3] Multivariate near infrared spectroscopy models for predicting the iodine value, CFPP, kinematic viscosity at 40 degrees C and density at 15 degrees C of biodiesel	Espectroscopia de Infravermelho Próximo (NIR)
P. Adewale et al. [1] Determination of the iodine value and the free fatty acid content of waste animal fat blends using FT-NIR	Espectroscopia de Infravermelho Próximo (NIR)

M. Rezaee et al. [40] desenvolveram um modelo de RNA preditivo para estudar e identificar os fatores que influenciam o tamanho de partícula e viscosidade de sódio carregados de diclofenaco de semente de palma de óleo ésteres em nano emul-

sões. O tamanho de partícula e a viscosidade de amostras em várias composições preparadas de acordo com o tempo e taxa de cisalhamento elevado foram medidos como saída. O modelo desenvolvido representou um desempenho de alta qualidade de uma rede neural e a sua capacidade em modelagem e identificar os fatores críticos que controlam a preparação das nanoemulsões.

H. Sharon et al. [46] utilizou um modelo de RNA com o poder de freio e mistura de biocombustíveis como entrada e consumo de combustível freio específico, eficiência térmica do freio, Óxido de Nitrogênio (NO), Hidrocarbono, Monóxido de carbono e densidade como saída. O algoritmo backpropagation foi utilizado e os dados obtidos a partir de ensaios de motores foi utilizado para treinamento da rede. A rede neural treinada com um coeficiente de correlação de 0.9989 e 0.999 gerou melhores estimativas para diesel com 15% (B15), 30% (B30), 60%(B60) e 90%(B90) de biodiesel, os resultados foram considerados aceitáveis.

Além dos biocombustíveis, Y. Çay et al. [13] conduziu seu trabalho com redes neurais para predição de desempenho do motor e emissão de escape para metanol e gasolina. Para construção do modelo de predição foi utilizado quatro diferentes algoritmos de aprendizado. A pesquisa mostrou a habilidade das redes neurais em prever as emissões de desempenho do motor e escape com coeficiente de correlação de 0,99.

Podemos observar que as RNAs estão sendo aplicados com sucesso em diversos domínios correlatados aos combustíveis e biocombustíveis e que os resultados nos possibilitam utiliza-las, uma revisão é feita em Jahirul et al. [24] demonstrando que a modelagem de RNAs tem um alto potencial para contribuir para o desenvolvimento de sistemas de energias renováveis, acelerando a pesquisa biodiesel. A tabela nº 2.1 demonstra alguns dos trabalhos desenvolvidos que utilizam técnicas de RNAs e outros modelos para predição de propriedades de qualidade de combustíveis e biocombustíveis.

## **3 Abordagem de Desenvolvimento de RNAs**

### **Proposta para Predição de Propriedades de Biodiesel**

Neste capítulo será apresentada a metodologia de desenvolvimento no processo de treinamento das RNAs, os principais critérios de configuração das simulações e os conceitos envolvidos no projeto da solução proposta.

#### **3.1 Processo de construção de RNAs**

A construção de uma topologia de rede, considerada mais apropriada para resolver um problema específico, é usualmente efetuada de forma empírica. Pois tal dimensionamento depende, entre outros fatores, do algoritmo de aprendizagem utilizado, da forma como as matrizes de pesos foram iniciadas, da complexidade do problema a ser mapeado, da disposição espacial das amostras e, por sua vez, da qualidade do conjunto de treinamento disponível [15]. Além disso, deve-se levar em consideração a quantidade de ruídos existentes nos dados utilizados.

Em termos práticos, no processo de construção de RNAs pode-se identificar duas fases principais, fase de treinamento e fase de operação, cada fase possui as seguintes atividades:

##### **I Fase de Treinamento:**

- 1 Selecionar padrões de treinamento e teste;
- 2 Pré-processar padrões (normalizar os dados);
- 3 Definir topologias candidatas;
- 4 Treinar topologias candidatas;
- 5 Escolher melhor topologia;

##### **II Fase de Operação:**

- 1 Apresentar padrões de entrada;
- 2 Pré-processar padrões (normalizar dados);
- 3 Aplicar padrões nas entradas da rede treinada;
- 4 Obter resultados das saídas da rede treinada;
- 5 Pós-processar resultados (desnormalizar dados);

Essas atividades implicam no aspecto de se encontrar as melhores topologias candidatas para a resolução do problema em questão. Assim, cada atividade ajuda a melhorar o treinamento e encontrar aquela rede que ofereça um melhor desempenho e eficiência computacional. No decorrer do capítulo serão descritas cada fase e as configurações de elaboração da metodologia de treinamento e testes das redes.

### 3.1.1 Pré-treinamento e ferramenta utilizada

No processo de construção e otimização das RNAs foi utilizada a ferramenta *Matlab Version 2013b Neural Network Toolbox*. *Neural Network Toolbox* permite a criação, treinamento e simulação de redes neurais para a modelagem de sistemas não-lineares complexos que não são facilmente modelados com uma equação linear. Permitem o apoio a aprendizagem supervisionada com *feedforward*, redes com funções de base radial, e redes dinâmicas. Ele também suporta o aprendizado não supervisionado com mapas de auto-organização e redes competitivas [5].

Primeiramente, foi analisada e selecionada a estrutura do tipo Perceptron Multiplas Camadas (MLP) para a resolução do problema, devido ser uma das populares e aplicada com sucesso em diversos domínios [31]. Em seguida, foram delimitados os algoritmos de convergência de redes neurais para serem testados:

- Backpropagation padrão;
- Backpropagation com *momentum*;
- Backpropagation BFGS Quase-Newton;
- Levenberg-Marquardt;
- *Resilient-propagation*;

- Gradiente Conjugado Escalonado;
- *One-Step-Secant*;

Os algoritmos selecionados são divididos em duas categorias de estratégia de treinamento, sendo métodos de primeira ordem e segunda ordem. Os métodos de primeira ordem são aqueles no qual somente o gradiente local determina a direção da minimização, conhecidos como métodos de gradiente descendente. Os métodos de segunda ordem são considerados mais eficientes para treinar redes neurais do tipo MLP [47].

Os algoritmos de segunda ordem resultam de um rigor matemático baseado em modelos de otimização não - linear irrestrita, não apresentando uma ligação natural com a inspiração biológica inicialmente proposta pelas RNAs. O método de Newton, por exemplo, é considerado como o método básico local que usa informações de segunda ordem, a sua aplicação requer o cálculo da matriz Hessiana [47].

Depois de selecionada a estrutura e tipo de treinamento a ser aplicado na resolução do problema pode-se passar para as fases de treinamento e operação propriamente dita. Estas fases podem ocorrer basicamente em paralelo por meio de sub-rotinas (funções e procedimentos), cujas instruções internas podem também ser refinadas ou subgrupadas para o propósito de elevação da eficiência computacional. [15]

### 3.1.2 Dados experimentais

Neste estudo, os dados utilizados para treinamento, validação e teste das redes foram reunidos da literatura [12, 16, 22, 27, 32, 33, 35, 43, 45] e outros cedidos pelo Laboratório de Análise e Pesquisa em Química Analítica de Petróleo e Biocombustíveis (LAPQAP/UFMA). Um conjunto de 98 amostras com 13 ésteres de ácidos graxos (EAGs) de entrada e na saída os parâmetros de qualidade índice de iodo e viscosidade cinemática.

Pode-se verificar pela Figura 3.1 a distribuição de cada tipo de éster de ácido graxos nas amostras de biodiesel coletadas e as predominâncias dos ésteres, como o ácido oleico (C18:1), ácido linoleico (C18:2), ácido láurico (C12:0) e ácido palmítico (C16:0). A descrição para os EAGs neste trabalho (XX:X) é sobre a informação

do número de carbonos (XX) e o número a direita (X) representa o número de insaturações da molécula. As amostras de biodiesel são derivadas de vários tipos de óleos e oleogênicas, entre elas, biodiesel metílico e etílico de soja e mamona, dendê, linhaça, colza, pinhão manso, moringa, babaçu, gordura de aves e algas.

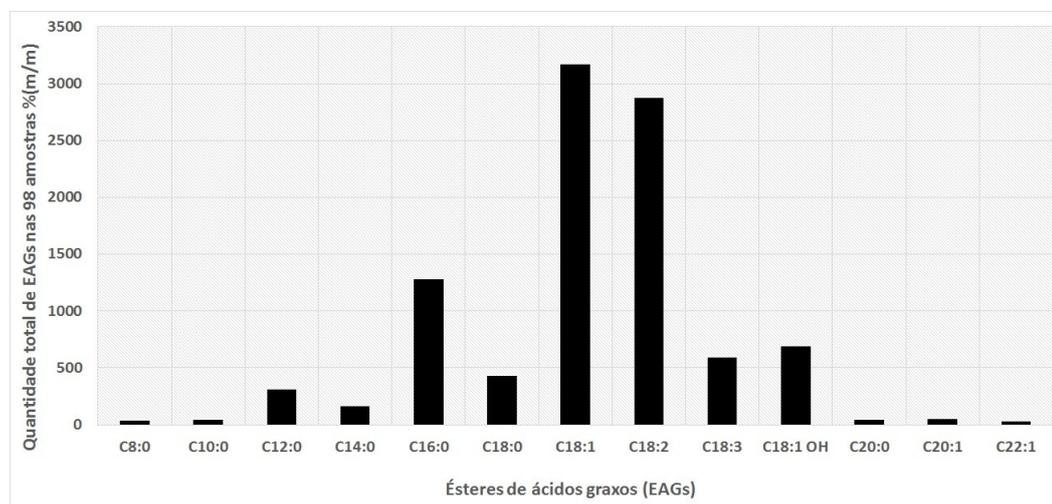


Figura 3.1: Quantidade de ésteres predominantes nas amostras de biodiesel

### 3.1.3 Pré-processamento dos dados

Técnicas de pré-processamento são usualmente aplicadas nos dados antes do treinamento como modo de acelerar a convergência. Uma delas é a normalização dos dados dentro de um intervalo uniforme, isso é essencial para evitar que um número maior possa vir a substituir os mais pequenos, além de evitar a saturação prematura de nós escondidos, o que impede o processo de aprendizagem. Isto é especialmente verdadeiro quando os dados de entrada reais tomam grandes valores [4].

Não há um procedimento padrão para normalizar as entradas e saídas dos dados. Porém, é preciso levar em consideração que o escalonamento das amostras seja feito dentro da faixa de valores dinâmica das funções de ativação das camadas escondidas, tipicamente representadas pelas funções logística ou tangente hiperbólica, com intuito de evitar a saturação dos neurônios [15].

Entre as técnicas mais utilizadas pode-se citar a normalização baseada no princípio dos segmentos proporcionais (teorema de Tales) ilustrado na Equação 3.1, em que um conjunto de valores definidos inicialmente num intervalo fechado entre  $x^{min}$

e  $x^{max}$ , será convertido para um domínio proporcional entre -1 e 1, o qual pode estar representando as próprias faixas de variações das funções de ativação.

$$y = (y^{max} - y^{min}) * (x - x^{min}) / (x^{max} - x^{min}) + y^{min} \quad (3.1)$$

Como recomendado, os padrões utilizados no presente trabalho foram pré-processados entre -1 e 1 de acordo com o teorema de Tales e de forma similar também, realizou-se instruções de pós-processamento na fase de teste da rede.

## 3.2 Aspectos de especificação topológica das redes

A partir do momento que a estrutura da rede, os padrões de treinamento e teste foram selecionados e pré-processados, pode-se passar para fase de configuração topológica da rede. Em suma, serão levados em consideração o número de camadas ocultas da rede, a quantidade de neurônios, as funções de ativação utilizadas, as técnicas de validação, o número de treinamento, as técnicas de parada e os aspectos de *overfitting* e *underfitting*.

### 3.2.1 Definição de topologias candidatas

As RNAs simuladas foram configuradas de duas maneiras: RNAs com uma camada oculta e RNAs com duas camadas ocultas. Esta decisão foi tomada partindo do princípio que um perceptron com multiplas camadas treinado com o algoritmo *backpropagation* pode ser visto como um veículo prático para realizar um mapeamento não-linear de entrada-saída de natureza geral [21].

Assim, pelo teorema da aproximação universal, uma única camada oculta aproxima qualquer função contínua, o problema com uma única camada é que os neurônios tendem a interagir globalmente entre si, que, em situações complexas essa interação torna difícil de melhorar a aproximação de um ponto sem piorá-la em algum outro ponto. Por outro lado, com duas camadas o processo de aproximação se torna mais gerenciável, ou seja, aproxima qualquer função matemática [21] [47]. Desse modo, simulações com uma e duas camadas ocultas foram necessárias para analisar qual melhor se adaptava na resolução da problemática.

Além do número de camadas ocultas, a escolha de uma quantidade ótima de neurônios nas camadas ocultas é uma das tarefas mais críticas na concepção de uma RNA. Pois uma rede com poucos nós nas camadas ocultas seriam incapazes de diferenciar padrões complexos, por outro lado, se uma rede tem muitos neurônios, vai absorver ruídos dos dados de entrada levando a má generalização dos dados [4]. No presente trabalho os neurônios variaram em cada treinamento de 2 a 25 neurônios em cada camada.

Existem outros fatores que são importantes no processo de construção da rede, como a função de ativação. O cálculo de um neurônio MLP no processo de treinamento requer o conhecimento da derivada da função de ativação associada aquele neurônio, e para esta derivada existir a função tem que ser contínua [21]. Funções que satisfazem essa condição foram utilizadas no presente trabalho, a função logística e a função tangente hiperbólica:

$$\text{logsig}(n) = \frac{1}{1 + e^{-u \cdot n}} \quad (3.2)$$

$$\text{tansig}(n) = \frac{1 - e^{-u \cdot n}}{1 + e^{-u \cdot n}} \quad (3.3)$$

### 3.2.2 Seleção de melhores topologias candidatas

Uma das técnicas mais utilizadas para seleção das melhores topologias candidatas é a validação cruzada (*cross-validation*), cujo propósito é avaliar a aptidão de cada uma quando aplicada a um conjunto de dados que seja diferente daquele usado no ajuste de seus parâmetros livres [15].

Os três métodos geralmente empregados são validação cruzada por amostragem aleatória (*radom subsampling cross-validation*), validação cruzada em  $k$ -partições (*k-fold cross validation*) e validação cruzada por unidade (*leave-one-out cross-validation*). No presente trabalho dois métodos de validação cruzada foram utilizados. Nas simulações iniciais foi feito uso da validação cruzada por amostragem aleatória, em que o conjunto total de amostras disponíveis é aleatoriamente dividido em duas partes, isto é, subconjunto de treinamento e subconjunto de teste. Nesta sistemática a quantidade de ensaios das topologias candidatas deve ser repetida várias vezes para que haja a

contemplanção de amostras diferentes tanto no subconjunto de treinamento como no de teste, e o desempenho médio global seja compilado [15].

Entretanto, em vista da quantidade de dados disponíveis para simulação e resultados preliminares obtidos pelo método de validação por amostragem aleatória não terem sido satisfatórios, optou-se por utilizar o método de validação de  $k$ -partições. Neste método as  $(k - 1)$  partições serão usadas para treinamento, ao passo que a partição restante ficará para teste, aqui o processo se repete  $k$  vezes até que todas as partições tenham sido utilizadas como conjunto de teste [29].

Além do critério de validação cruzada  $k$ -partições com  $k = 11$  como escolha da melhor rede entre as candidatas, a técnica de treinamento com parada antecipada (*early - stopping*) foi utilizada com o objetivo de detectar quando *overfitting* e/ou *underfitting* começar durante o treinamento de uma rede [37]. Este processo implica em que cada topologia candidata é constantemente checada pela aplicação do conjunto de teste, sendo finalizado quando começar a haver elevação do erro quadrático (frente aos subconjuntos de teste) entre épocas sucessivas [15].

*Overfitting* ocorre enquanto o erro no conjunto de treinamento diminuiu, porém em algum momento, o erro de validação começa a subir. Se o treinamento não parar neste ponto mínimo, a rede aprenderá após este ponto essencialmente ruídos contidos nos dados de treinamento [21] [37]. Em contrapartida, o *underfitting* se configura quando uma topologia com número muito reduzido de neurônios pode ser insuficiente para a extração e armazenamento de características que permitem à rede implementar as hipóteses a respeito do comportamento do processo. Neste caso, o erro quadrático tanto na fase de treinamento como na fase de testes é bem significativo [15].

A definição da quantidade de verificações de erros para parada antecipada foi obtida mediante testes das redes, com os seguintes valores de erros: 6, 10, 15, 20 e 25. Usou-se também o algoritmo de Levenberg Marquardt e função de ativação tangente hiperbólica. O cálculo do erro usado para mensurar o desempenho da rede foi o Erro Médio Quadrático ou *Mean Squared Error* (MSE), que interpreta a diferença entre o valor produzido pela rede, indicado por  $\hat{y}_i$ , e o valor real,  $y_i$  [31]:

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (3.4)$$

Desse modo, 10  $k$ -partes foram usadas para treinamento e validação; e a partição  $k=11$  usada para teste. As entradas e saídas de dados para treinamento, validação e teste foram pré-processadas entre -1 e 1. A Regressão Linear (*Linear Regression* -  $R$ ) e a Raiz Quadrada Média do Erro (*Root Mean Square Error* -  $RMSE$ ) foram usadas para análise de desempenho. Os valores de  $R$  e  $RMSE$  são uma indicação de correlação entre as saídas previstas e as saídas reais da rede.

$$R = \alpha + \beta X_i + \epsilon_i \quad (3.5)$$

$$RMSE = \frac{\sqrt{\sum_{i=1}^n (\hat{y}_i - y_i)^2}}{n} \quad (3.6)$$

Ainda em relação aos critérios de configuração das topologias de rede, os pesos da rede foram iniciados aleatoriamente, fixou-se um limite de 10000 mil épocas (quantidade de ciclos) de treinamento para cada topologia e uma taxa de aprendizagem no valor de 0.01. No final das simulações as 10 melhores redes foram salvas. Pode-se visualizar pela Tabela 3.1 um resumo dos principais parâmetros de configuração simulados.

**Tabela 3.1:** Configuração das topologias de redes simuladas.

Parâmetros de treinamento	Valor
Algoritmos	Backpropagation padrão Backpropagation com <i>momentum</i> e taxa de aprendizagem BFGS Quase-Newton Levenberg-Marquardt Resilient-propagation One-Step-Secant Gradiente Conjugado Escalonado
Função de ativação	Função tansig Função logsig
Número de camadas ocultas	1 e 2
Função de desempenho	MSE
Critério de generalização	Validação cruzada em $k$ -partições e parada antecipada
Valor máximo de verificações de parada antecipada	6, 10, 15, 20 e 25
Número máximo de épocas	10000
Taxa de aprendizagem	0.01
Número de neurônios por camada	2 - 25

## 4 Resultados e Discussões

Neste capítulo serão apresentados os resultados obtidos das simulações das RNAs na predição de índice de iodo e viscosidade cinemática, testadas tanto com uma e duas camadas ocultas. Além de demonstrar o porquê do critério de escolha do valor de parada antecipada para obtenção de uma melhor arquitetura de rede na predição dessas propriedades.

### 4.1 Considerações sobre o critério de parada antecipada

Neste estudo foram utilizadas 98 amostras de ésteres de ácidos graxos para predição tanto de índice de iodo, como de viscosidade cinemática. Foram desenvolvidas e testadas mais de 576 combinações para cada algoritmo para se obter as melhores topologias candidatas para predição de propriedades de qualidade do biodiesel. Vale ressaltar que o uso da validação cruzada com parada antecipada é um importante passo para evitar *overfitting* e *underfitting*. Assim, a definição do número de verificações de erro é um ponto essencial e crítico no treinamento da rede.

A definição do valor de erros para parada antecipada foi obtida por meio de simulações, com valores de 6, 10, 15, 20 e 25. O Levenberg-Marquardt foi o algoritmo de convergência usado para teste dos valores de erros, com duas camadas ocultas e função de ativação tangente hiperbólica para averiguar qual valor produzia melhor performance na predição da viscosidade e índice de iodo. As redes foram treinadas com separação de dados por amostra aleatória e por  $k$ -partições.

Os resultados dos coeficientes de correlação obtidos das simulações para escolha do valor de parada antecipada é reportado na Tabela 4.1 para predição de índice de iodo e na Tabela 4.2 para a viscosidade cinemática. Dos resultados obtidos pode-se aferir que valores de parada antecipada muito baixos ou valores muito altos não permitem que a rede aprenda o suficiente para generalizar bem os dados de treinamento, validação e teste; assim, classificando incorretamente novos dados.

A Tabela 4.1 mostra que os resultados apresentados para valores de parada antecipada iguais a 6, por exemplo, geraram coeficientes de correlação de 0.98 para treinamento, 0.99 para validação e 0.35 para teste, demonstrando que a rede não conseguiu generalizar bem dados que não estavam presentes no treinamento. E para valores de parada antecipada iguais a 25, os resultados mostrados também não conseguiram prever corretamente os padrões gerando coeficientes de correlação de 0.98 para treinamento, 0.99 para validação e 0.54 para teste, em uma arquitetura de rede de 10 neurônios na primeira camada oculta e 15 na segunda.

Também pode-se ver que as redes que generalizaram e obtiveram alto índice de correlação para predição de índice de iodo foram os valores de 20 erros de verificação para parada antecipada, como mostra a rede de arquitetura 12-5, que indica 12 neurônios na primeira camada e 5 neurônios na segunda, com valores de predição de 0.97 para treinamento, 0.92 para validação e 0.86 para teste. Com base nestes resultados obtidos com valor de erro igual a 20, este valor foi usado para simulações com os outros algoritmos de treinamento para prever o índice de iodo.

**Tabela 4.1:** Resultados dos testes para escolha do valor de parada antecipada para índice de iodo.

Configuração da Rede	Valor de parada antecipada	Treinamento - R	Validação - R	Teste - R	Separação dos dados
8-15	6	0.98	0.99	0.35	Aleatória
7-7	6	0.96	0.95	0.72	k-partições
6-5	10	0.72	0.97	0.82	Aleatória
6-21	10	0.98	0.98	0.78	k-partições
7-10	15	0.97	0.99	0.88	Aleatória
11-5	15	0.94	0.99	0.70	k-partições
<b>8-4</b>	<b>20</b>	<b>0.96</b>	<b>0.97</b>	<b>0.93</b>	<b>Aleatória</b>
<b>12-5</b>	<b>20</b>	<b>0.97</b>	<b>0.92</b>	<b>0.86</b>	<b>k-partições</b>
5-9	25	0.56	0.88	0.95	Aleatória
10-15	25	0.98	0.99	0.54	k-partições

A escolha do valor de erro de parada antecipada da viscosidade é mostrada na Tabela 4.2. Aqui pode-se ver que os coeficientes de predição que obtiveram os maiores índices foram aqueles com valor de parada antecipada igual a 10. A arquitetura de rede 11-5 apresentou coeficientes de 0.99, 0.99 e 0.96 para treinamento, validação e teste respectivamente. Como também na arquitetura 5-3 que apresentou valores de 0.98, 0.99 e 0.99 para treinamento, validação e teste respectivamente. Valores menores que 10 ou maiores que esse valor não conseguiram obter coeficientes de correlação tão

altos. Desse modo, o valor de parada antecipada escolhido para as simulações com os outros algoritmos foi de 10 para a propriedade viscosidade.

**Tabela 4.2:** Resultados dos testes para escolha do valor de parada antecipada para viscosidade.

Configuração da Rede	Valor de parada antecipada	Treinamento - R	Validação - R	Teste - R	Separação dos dados
12-4	6	0.99	0.88	0.97	Aleatória
5-10	6	0.99	0.86	0.97	k-partições
<b>11-5</b>	<b>10</b>	<b>0.99</b>	<b>0.99</b>	<b>0.96</b>	<b>Aleatória</b>
<b>5-3</b>	<b>10</b>	<b>0.98</b>	<b>0.99</b>	<b>0.99</b>	<b>k-partições</b>
4-10	15	0.99	0.98	0.70	Aleatória
8-12	15	0.99	0.99	0.80	k-partições
3-18	20	0.99	0.99	0.83	Aleatória
4-5	20	0.99	0.99	0.45	k-partições
13-2	25	0.98	0.99	0.83	Aleatória
6-7	25	0.99	0.99	0.73	k-partições

## 4.2 Resultados obtidos para índice de iodo

Como relatado no Capítulo 2, o índice de iodo indica o grau de insaturação do óleo ou gordura, considerando que o iodo reage com as duplas ligações; verifica-se que quanto maior o grau de insaturação (duplas ligações), maior será o índice de iodo. Esta propriedade está relacionada com fatores de qualidade do biodiesel, tais como a polimerização, a estabilidade e a viscosidade [20]. Assim é uma importante propriedade que análise de qualidade de biodiesel. Aqui serão apresentados os resultados obtidos para predição de índice de iodo por meio de RNAs.

### 4.2.1 Resultados obtidos com uma camada

Nas simulações realizadas com RNAs de apenas uma camada para predição de índice de iodo, pode-se observar os melhores resultados na Tabela 4.3, na qual são mostrados os coeficientes de correlação e o valor de RMSE para cada algoritmo de treinamento testado. O algoritmo *backpropagation* padrão foi o que obteve índices mais baixos para predição. Para a rede com 11 neurônios na camada oculta, os coeficientes de correlação foram de 0.53, 0,78 e 0,79 no treinamento, validação e teste. O RMSE também demonstrou que o algoritmo padrão não oferece uma boa predição, em vista de o erro médio aparecer com valores altos.

O algoritmo que demonstrou melhor desempenho na predição de índice de iodo para uma camada foi o *Resilient-Propagation* que chegou a um coeficiente de correlação de 1 na fase de teste, com um RMSE de 0.0917, e os coeficientes de correlação de 0.96 tanto para treinamento como para teste. Esta arquitetura de rede tem 19 neurônios em sua camada oculta com função de ativação logística. Os gráficos das análises de predição para índice de iodo de cada algoritmo com as melhores arquiteturas para uma camada, são apresentados pelas Figuras 4.1 a 4.7.

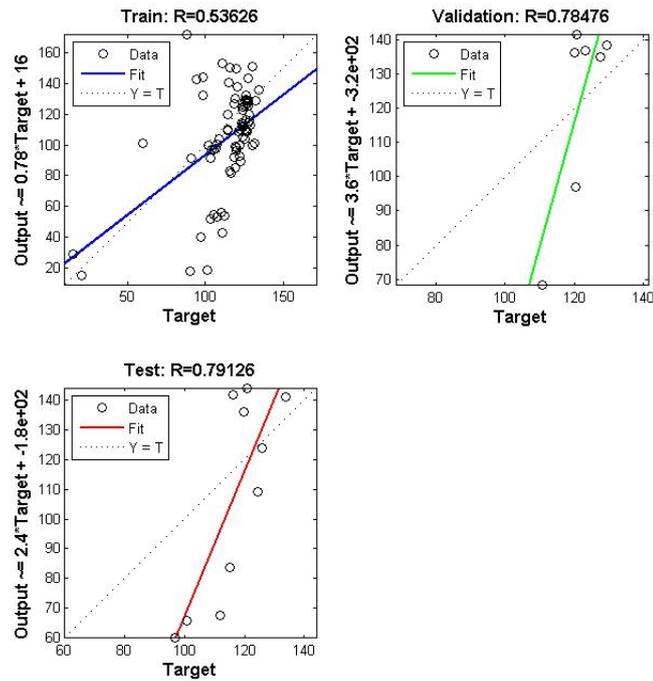
**Tabela 4.3:** Coeficientes de predição para índice de iodo com redes de uma camada oculta.

Configuração da Rede	Algoritmos de Treinamento	Função de Ativação	Treinamento R	RMSE	Validação R	RMSE	Teste R	RMSE
11	Backpropagation padrão	logsig	0.53	28.67	0.78	19.06	0.79	22.56
12	Backpropagation com <i>momentum</i> e taxa de aprendizagem	tansig	0.92	12.06	0.99	4.663	0.98	6.608
12	BFGS Quase-Newton	logsig	0.96	8.76	0.99	6.402	0.8743	14.57
23	Gradiente Conjugado Escalonado	logsig	0.95	9.85	0.95	7.873	0.93	12.2
19	Resilient-propagation	logsig	0.96	8.88	0.96	9.782	1	0.0917
10	One-Step-Secant	logsig	0.93	11.12	0.96	11.91	0.98	5.566
7	Levenberg-Marquardt	logsig	0.96	8.736	0.99	5.503	0.86	15.41

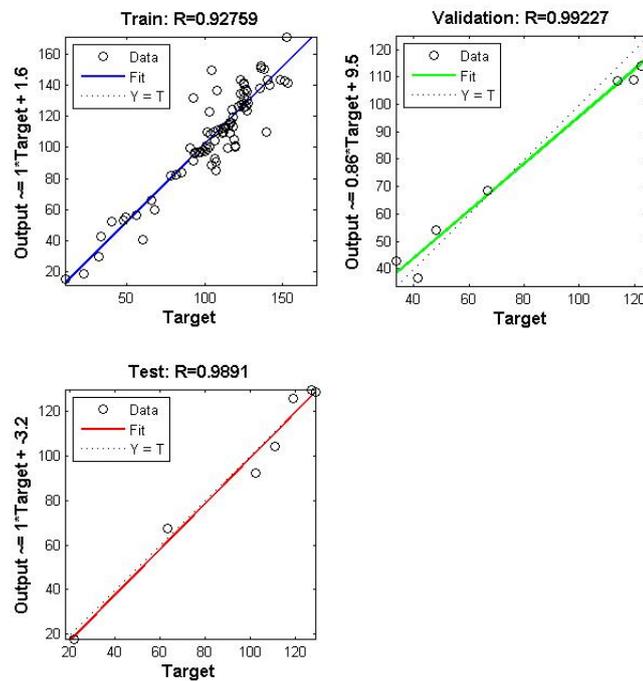
## 4.2.2 Resultados obtidos com duas camadas

A partir da simulação de redes com duas camadas, os coeficiente de correlação de predição de maior valor foram apresentados pelo algoritmo Levenberg-Marquardt, *Resilient-Propagation* e BFGS. A arquitetura da rede obtida com coeficiente de correlação de 0.99 para o treinamento, validação e teste é a rede 9-7 com o algoritmo de Levenberg-Marquardt e função de activação logsig. A Tabela 4.4 mostra o resultados obtidos de cada algoritmo que foi treinado com os seus respectivos coeficientes de correlação e RMSEs.

Como esperado, as redes testadas com apenas backpropagation padrão tiveram resultados inferiores quando comparadas com outros algoritmos testados devido a que ficou preso em mínimos locais, com  $R = 0.79$ ,  $R = 0.79$  e  $R = 0.32$ , respectivamente, para o treinamento, validação e teste numa arquitetura de rede 24-3. No entanto, os re-



**Figura 4.1:** Gráfico do coeficiente de correlação do índice de iodo para backpropagation padrão com uma camada.



**Figura 4.2:** Gráfico do coeficiente de correlação do índice de iodo para backpropagation com *momentum* com uma camada.

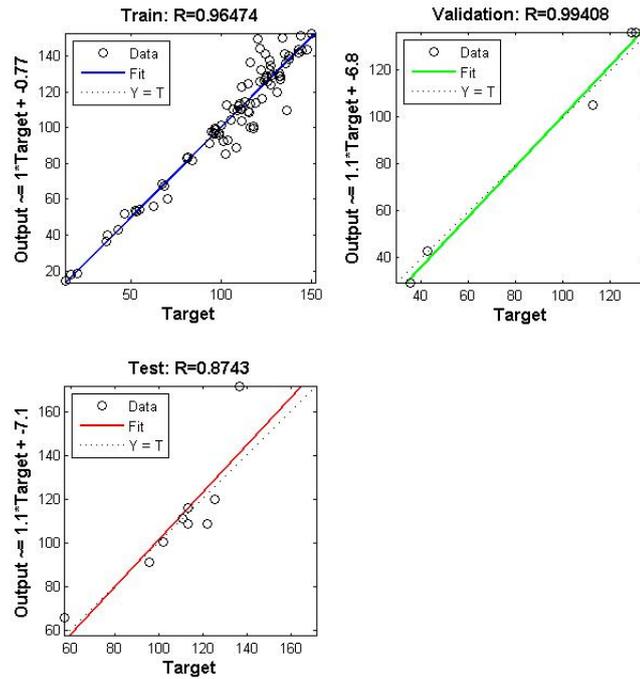


Figura 4.3: Gráfico do coeficiente de correlação do índice de iodo para BFGS com uma camada

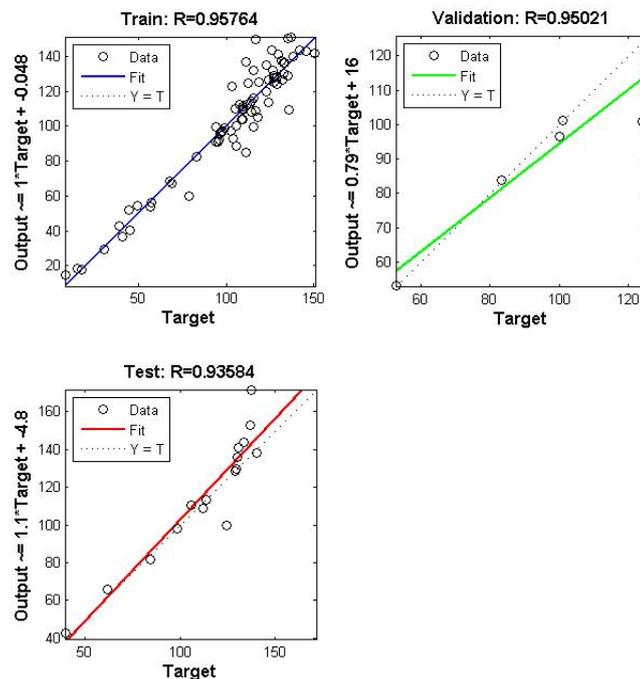


Figura 4.4: Gráfico do coeficiente de correlação do índice de iodo para gradiente conjugado escalonado com uma camada.

sultados obtidos por algoritmos Gradiente Conjugado Escalonado e One-Step-Secant permaneceram equivalentes.

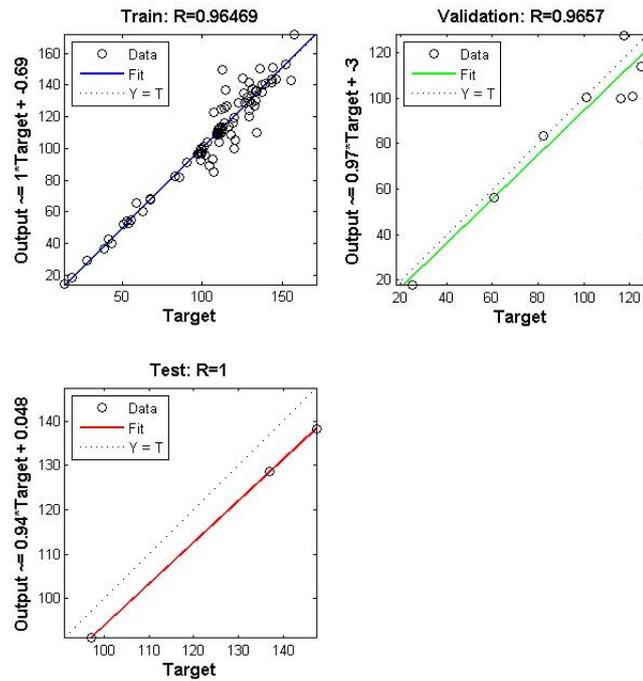


Figura 4.5: Gráfico do coeficiente de correlação do índice de iodo para Resilient-Propagation com uma camada.

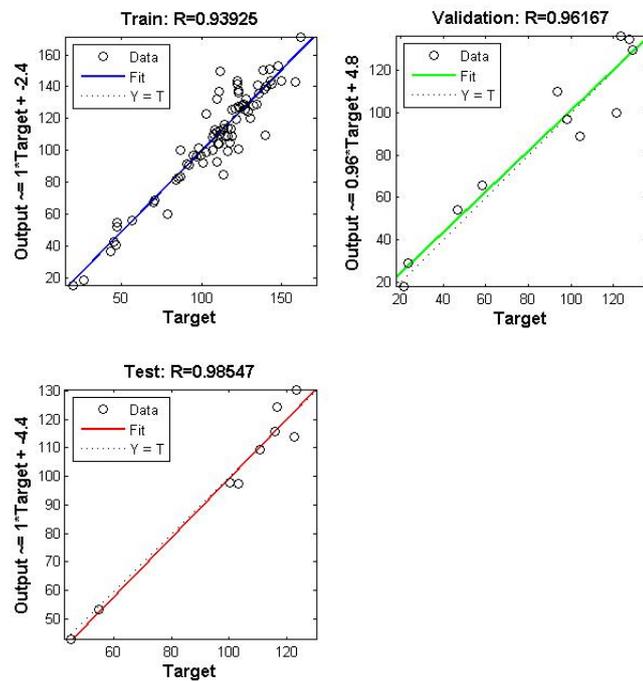
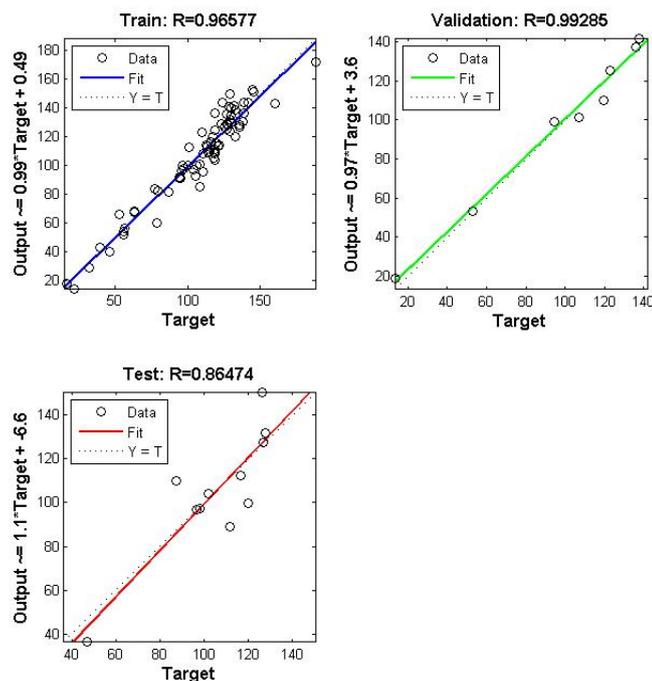


Figura 4.6: Gráfico do coeficiente de correlação do índice de iodo para One-Step-Secant com uma camada.



**Figura 4.7:** Gráfico do coeficiente de correlação do índice de iodo para Levenberg-Marquardt com uma camada.

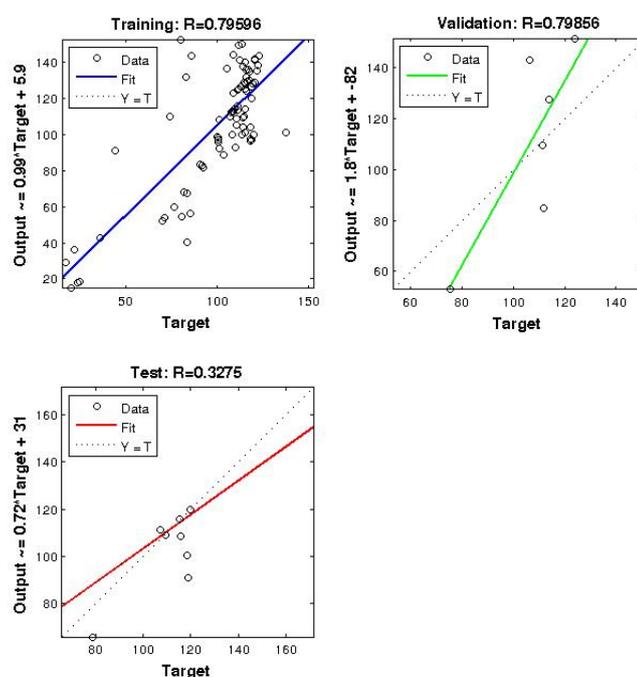
E, embora o tempo de execução de treinamento não ter sido medido no correspondente estudo, vale a pena destacar que alguns algoritmos têm mais tempo de treinamento, em vista do cálculo da matriz hessiana, tal como os algoritmos de BFGS e Levenberg Marquardt [47]. As Figuras de 4.8 a 4.14 mostram o coeficiente de correlação de valor obtido e o valor-alvo de previsão, resultante dos algoritmos de treinamento, demonstrando a capacidade da rede de duas camadas para generalizar e simular novos dados de éster de ácido graxo e prever o índice de iodo.

### 4.3 Resultados obtidos para viscosidade cinemática

A viscosidade é uma importante medida da resistência da vazão de um líquido associada à fricção ou atrito interno de uma parte do fluido que escoa sobre a outra, ocasionando formação de depósitos no motor. Quanto maior a viscosidade maior é a tendência do combustível causar problemas. Esta propriedade pode ser determinada por metodologias padronizadas, como ASTM D445 ou a ISO 3104. Por sua vez a análise da diferença da viscosidade entre o óleo de origem e o seu derivado aquíl éster pode ser utilizada para monitorar o biodiesel [28]. Nesta seção serão apresenta-

**Tabela 4.4:** Coeficientes de predição para índice de iodo com redes de duas camadas ocultas.

Configuração da Rede	Algoritmos de Treinamento	Função de Ativação	Treinamento R	RMSE	Validação R	RMSE	Teste R	RMSE
24-3	Backpropagation padrão	tansig	0.79	20.75	0.79	25.07	0.32	28.36
9-21	Backpropagation com <i>momentum</i> e taxa de aprendizagem	tansig	0.95	10.3	0.96	7.337	0.77	22.61
9-9	BFGS Quase-Newton	logsig	0.95	10.08	0.99	10.1	0.90	12.67
4-13	Gradiente Conjugado Escalonado	tansig	0.94	11.54	0.93	10.65	0.92	17.46
25-7	Resilient-propagation	tansig	0.97	8.008	0.97	4.445	0.97	12.5
12-4	One-Step-Secant	logsig	0.94	11.17	0.95	12.27	0.87	15.76
9-7	Levenberg-Marquardt	logsig	0.99	4.3035	0.99	2.677	0.99	2.5832

**Figura 4.8:** Gráfico do coeficiente de correlação do índice de iodo para backpropagation padrão com duas camadas.

dos os resultados e arquiteturas obtidas para predição da viscosidade utilizando RNAs com uma camada e com duas camadas.

### 4.3.1 Resultados obtidos com uma camada

Pode-se aferir dos resultados obtidos de predição de viscosidade com redes de uma camada que o algoritmo Levenberg-Marquardt mais uma vez obteve o melhor

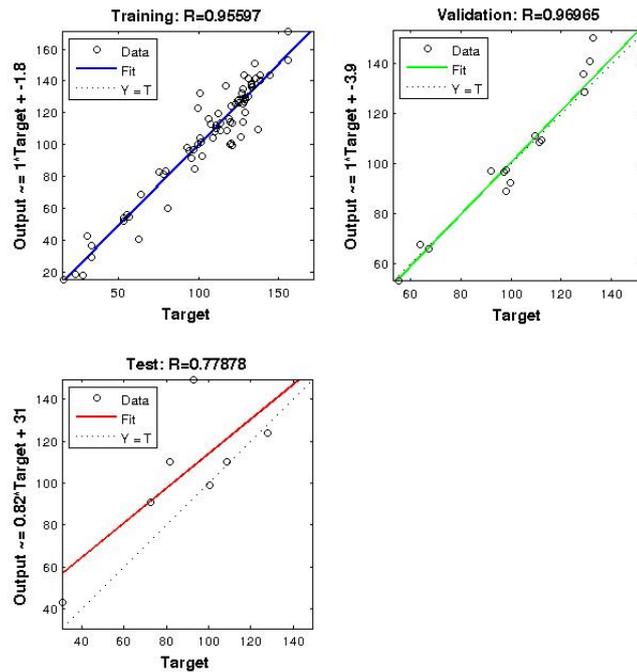


Figura 4.9: Gráfico do coeficiente de correlação do índice de iodo para backpropagation com *momentum* com duas camadas.

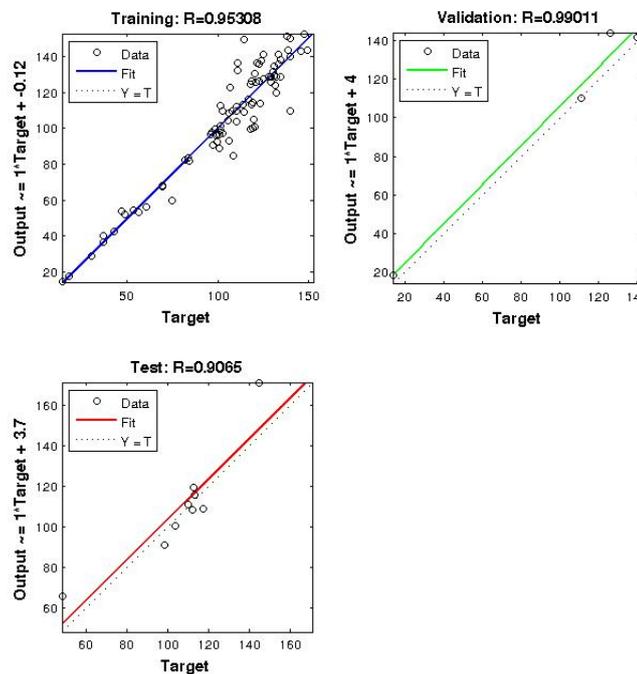
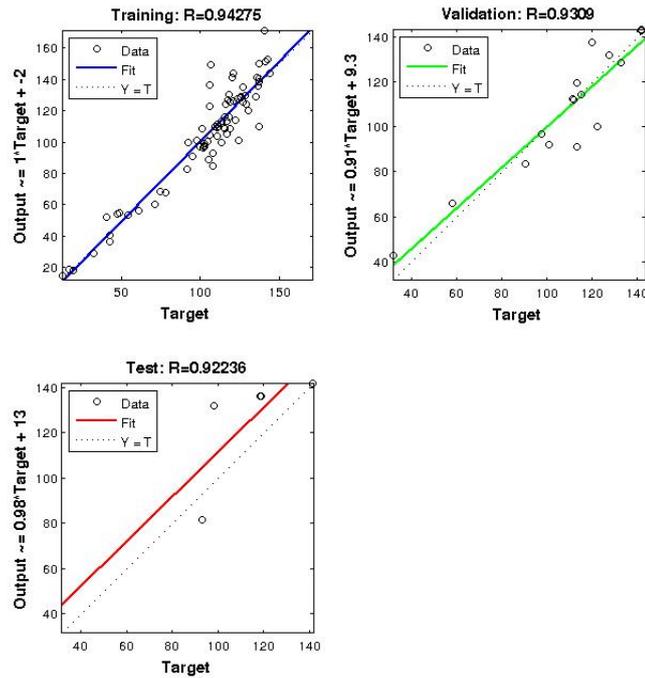
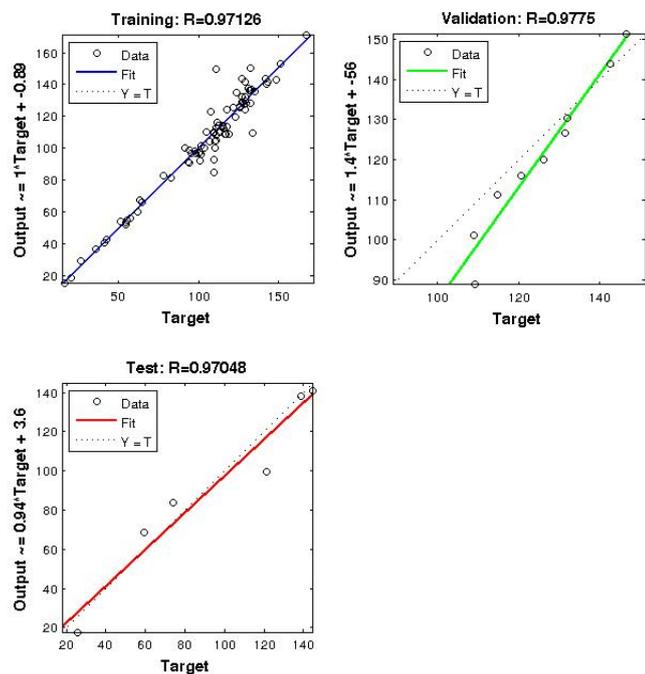


Figura 4.10: Gráfico do coeficiente de correlação do índice de iodo para BFGS com duas camadas.

coeficiente de correlação para rede com 8 neurônios na camada oculta. A regressão chegou a valores de 0.99 tanto para treinamento, validação e teste. O segundo melhor



**Figura 4.11:** Gráfico do coeficiente de correlação do índice de iodo para gradiente conjugado escalonado com duas camadas.



**Figura 4.12:** Gráfico do coeficiente de correlação do índice de iodo para Resilient-Propagation com duas camadas.

desempenho foi com o algoritmo BFGS que conseguiu generalizar com coeficientes de

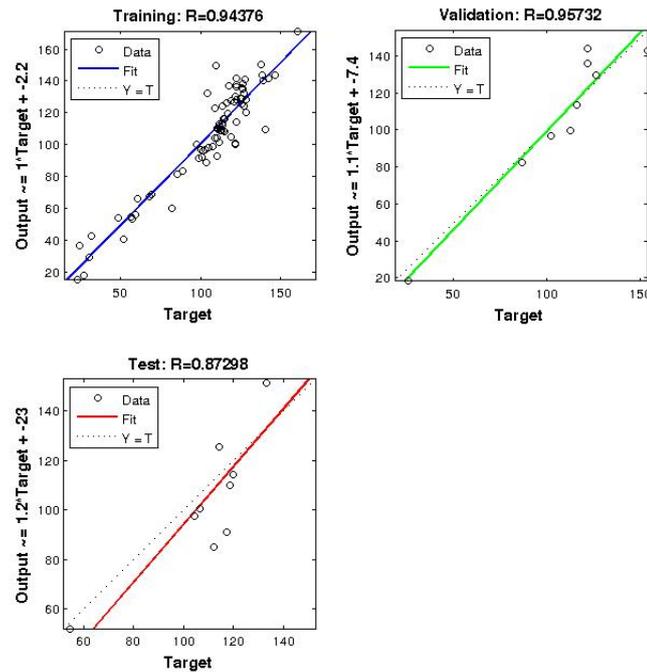


Figura 4.13: Gráfico do coeficiente de correlação do índice de iodo para One-Step-Secant com duas camadas.

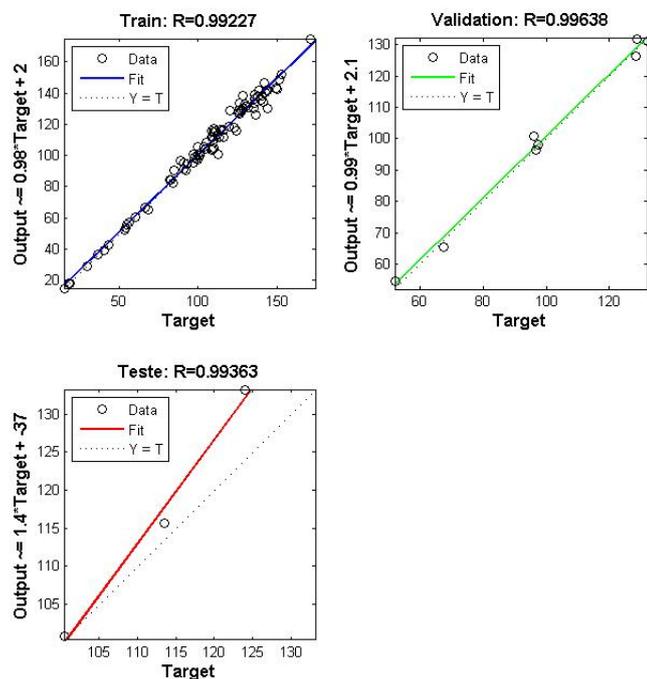


Figura 4.14: Gráfico do coeficiente de correlação do índice de iodo para Levenberg-Marquardt com duas camadas.

correlação de 0.99 para treinamento, 0.99 para validação e 0.96 para teste, que pode ser observado pela Tabela 4.5.

Além dos valores de regressão observados, vale ressaltar o RMSE obtido dos algoritmos de treinamento testados, diferentemente dos resultados obtidos pelo índice de iodo, todos os valores de RMSE ficaram abaixo de 1. Isso pode ser explicado pelo valor de parada antecipada que permitiu os baixos erros e também pelo limite estabelecido segundo as normas da ANP para qualidade da viscosidade na qual os dados utilizados para treinamento possuem. Como a norma da ANP para índice de iodo não estabelece um limite mínimo e máximo, essa propriedade apresenta características de predição diferente. As Figuras de 4.15 a 4.21 mostram o coeficiente de correlação das melhores redes de uma camada oculta para viscosidade.

**Tabela 4.5:** Coeficientes de predição para viscosidade com redes de uma camada oculta.

Configuração da Rede	Algoritmos de Treinamento	Função de Ativação	Treinamento R	RMSE	Validação R	RMSE	Teste R	RMSE
13	Backpropagation padrão	logsig	0.98	0.33	0.99	0.44	0.94	0.35
20	Backpropagation com <i>momentum</i> e taxa de aprendizagem	logsig	0.98	0.4305	0.96	0.2979	0.3332	0.6893
8	BFGS Quase-Newton	logsig	0.99	0.2791	0.99	0.3495	0.96	0.6685
9	Gradiente Conjugado Escalonado	logsig	0.99	0.33	0.77	0.4405	0.97	0.1911
22	Resilient-propagation	logsig	0.99	0.3186	0.41	0.5762	0.99	0.4966
17	One-Step-Secant	logsig	0.98	0.3895	0.97	0.1785	0.96	0.5556
8	Levenberg-Marquardt	logsig	0.99	0.276	0.99	0.2772	0.99	0.3803

### 4.3.2 Resultados obtidos com duas camadas

Dos resultados obtidos de previsão de viscosidade com duas camadas, observou-se que praticamente todos os algoritmos apresentaram um bom desempenho para generalização de dados novos. Exceto pelos algoritmos BFGS e *Resilient-Propagation* que obtiveram baixos desempenhos na fase de teste da rede, o que pode significar um *overfitting* do treinamento, em vista da quantidade de neurônios nas camadas ocultas, apesar da validação cruzada.

Pode-se comparar também o desempenho do *backpropagation* padrão, *backpropagation* com *momentum* e o Levenberg-Marquardt, os três obtiveram coeficientes de correlação de 0.99 para treinamento, validação e teste da rede. Porém o que obteve menor número de neurônios foi o Levenberg-Marquardt com uma arquitetura 5-6, ou

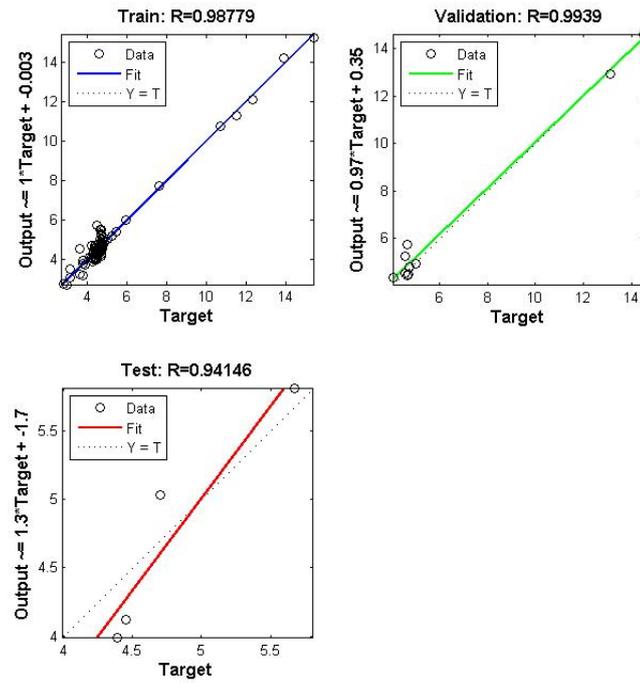


Figura 4.15: Gráfico do coeficiente de correlação da viscosidade para backpropagation padrão com uma camada.

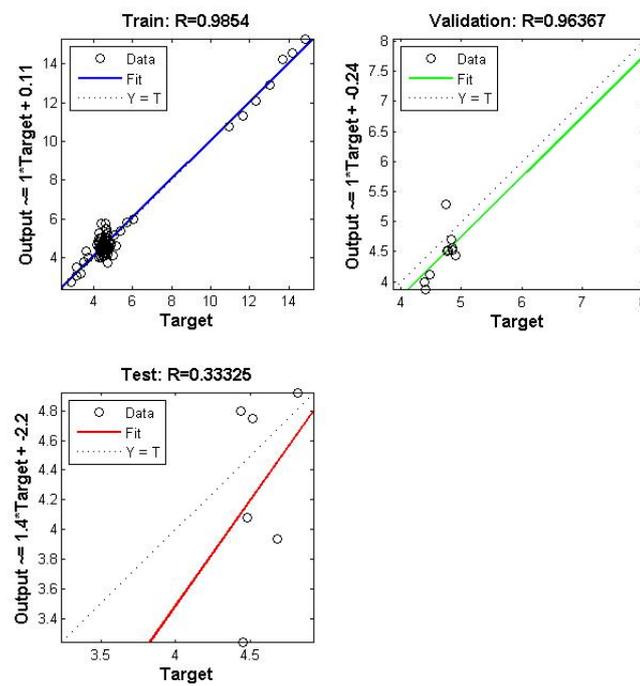


Figura 4.16: Gráfico do coeficiente de correlação da viscosidade backpropagation com *momentum* com uma camada.

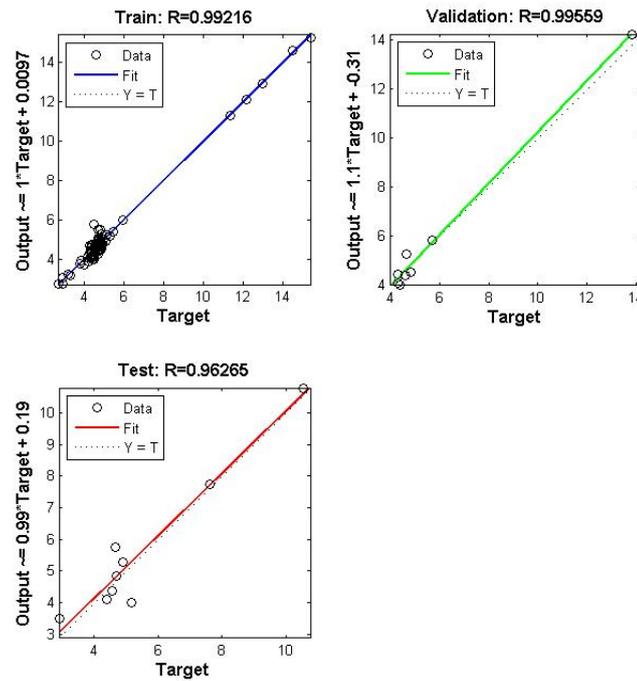


Figura 4.17: Gráfico do coeficiente de correlação da viscosidade para BFGS com uma camada.

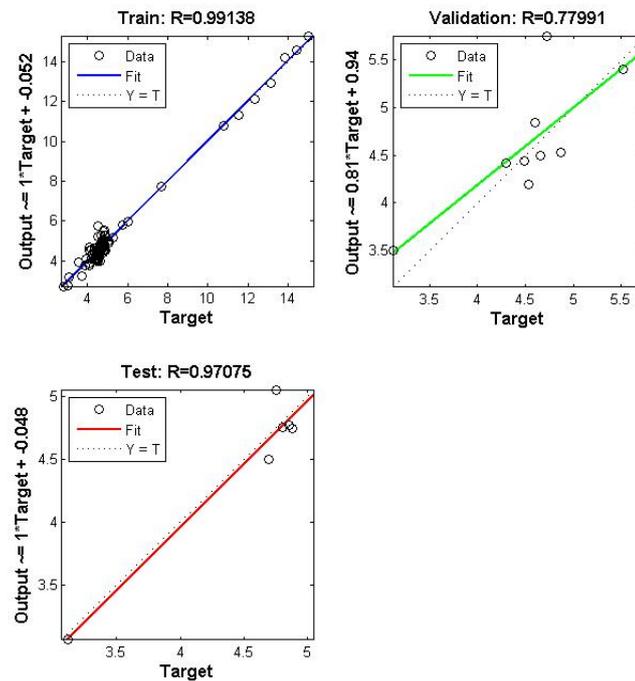


Figura 4.18: Gráfico do coeficiente de correlação da viscosidade para gradiente conjugado escalonado com uma camada.

seja, 5 neurônios da primeira camada oculta e 6 neurônios na segunda camada oculta. Enquanto que o *backpropagation* padrão e com *momentum* precisaram de um número

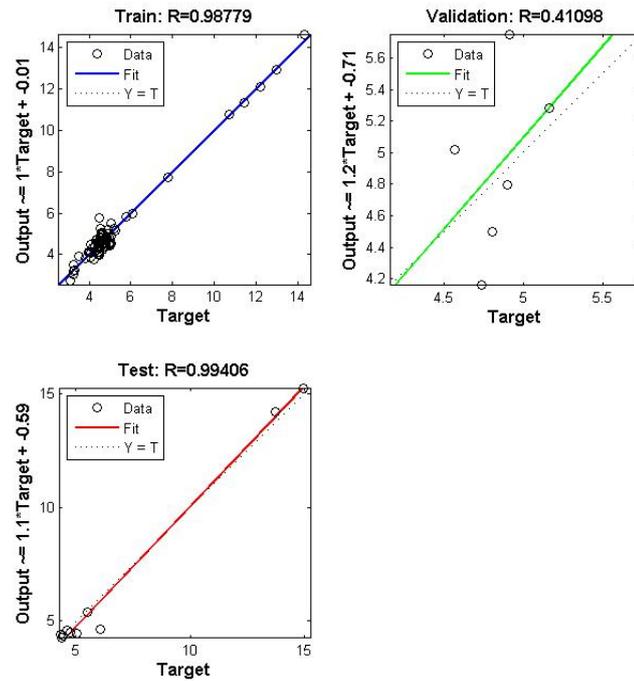


Figura 4.19: Gráfico do coeficiente de correlação da viscosidade para Resilient-Propagation com uma camada.

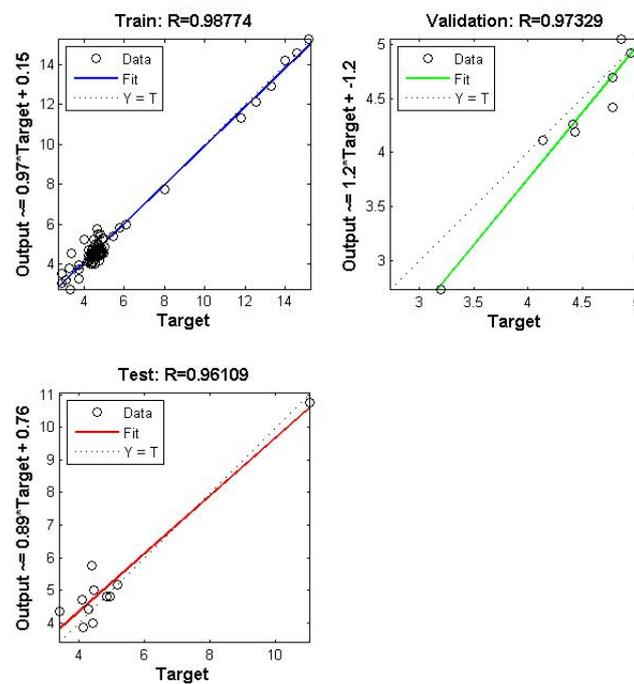
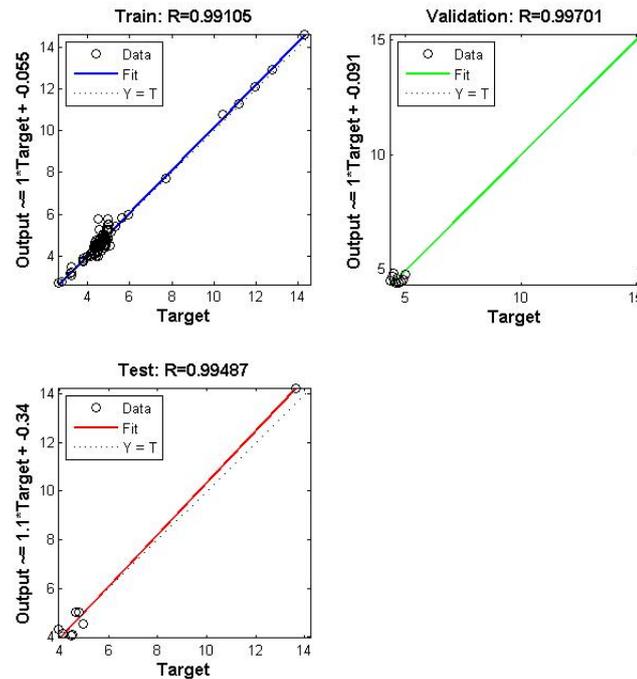


Figura 4.20: Gráfico do coeficiente de correlação da viscosidade para One-Step-Secant com uma camada.



**Figura 4.21:** Gráfico do coeficiente de correlação da viscosidade para Levenberg-Marquardt com uma camada.

superior de neurônios, com arquiteturas de 22-9 e 20-8, respectivamente, para generalizar bem com esses coeficientes.

A Tabela 4.6 mostra todos os coeficientes obtidos para cada algoritmo na predição da viscosidade. Trabalhos relatados [25,31,42] já demonstraram a habilidade do algoritmo Levenberg-Marquardt de predizer a propriedade viscosidade cinemática por meio de RNAs. Neste estudo confirma-se essa habilidade tanto para índice de iodo e viscosidade, enquanto que alguns trabalhos relacionados chegaram a apenas 0.97 na predição dessas propriedades, aqui obtive-se índices de 0.99 de predição.

É preciso considerar também que os resultados obtidos para viscosidade realizados com uma camada oculta, já é considerado suficiente de acordo com os testes, esses resultados obtidos de 0.99 de predição tanto para validação e teste do Levenberg-Marquardt, pode ser observado na Figura 4.21. Pois a finalidade do treinamento e técnicas de escolha de melhor topologia de rede, é encontrar a arquitetura que ofereça o melhor desempenho com menor custo computacional, ou seja, com menor quantidade de neurônios. As figuras de 4.22 a 4.28 mostram os melhores coeficientes de correlação obtidos com duas camadas ocultas para viscosidade.

Tabela 4.6: Coeficientes de predição para viscosidade com redes duas camadas ocultas.

Configuração da Rede	Algoritmos de Treinamento	Função de Ativação	Treinamento R	RMSE	Validação R	RMSE	Teste R	RMSE
22-9	Backpropagation padrão	logsig	0.98	0.3388	0.99	0.3721	0.99	0.2694
20-8	Backpropagation com <i>momentum</i> e taxa de aprendizagem	tansig	0.99	0.3159	0.99	0.5098	0.99	0.108
11-12	BFGS Quase-Newton	tansig	0.97	0.4619	0.99	0.2158	0.63	0.4493
18-13	Gradiente Conjugado Escalonado	tansig	0.99	0.3063	0.96	0.3698	0.98	0.4916
21-4	Resilient-propagation	logsig	0.98	0.316	0.99	0.418	0.32	0.3046
10-13	One-Step-Secant	logsig	0.98	0.342	0.94	0.2555	0.98	0.4765
5-6	Levenberg-Marquardt	tansig	0.99	0.1785	0.99	0.2252	0.99	0.418

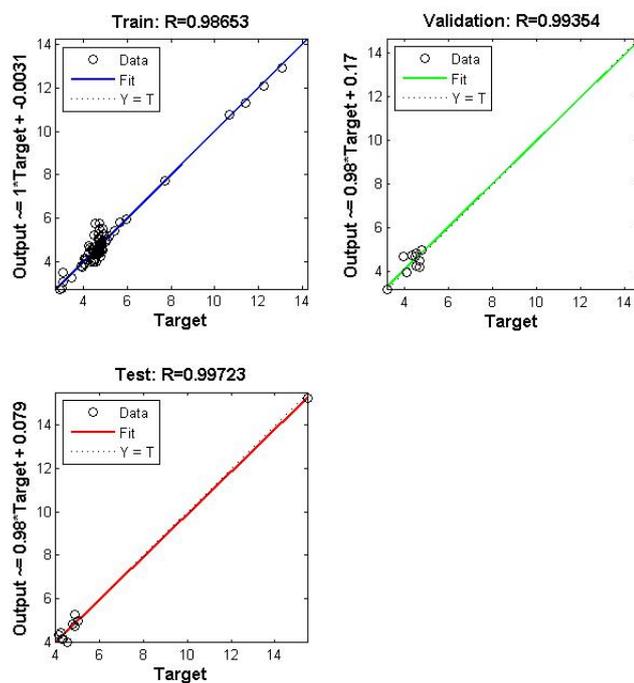


Figura 4.22: Gráfico do coeficiente de correlação da viscosidade para backpropagation padrão com duas camadas.

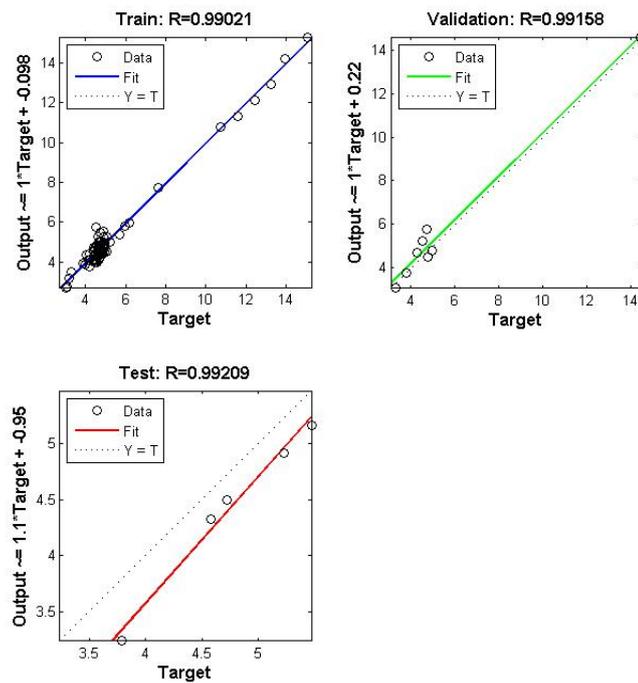


Figura 4.23: Gráfico do coeficiente de correlação da viscosidade backpropagation com *momentum* com duas camadas.

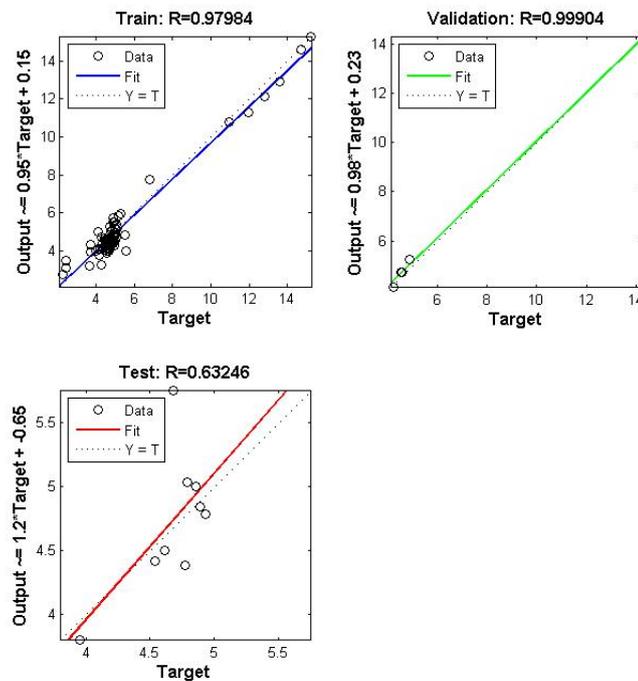


Figura 4.24: Gráfico do coeficiente de correlação da viscosidade para BFGS com duas camadas.

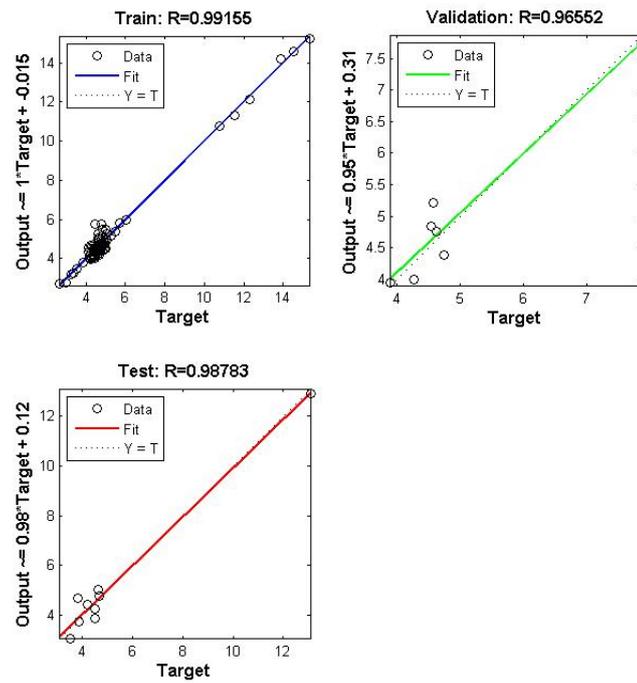


Figura 4.25: Gráfico do coeficiente de correlação da viscosidade para gradiente conjugado escalonado com duas camadas.

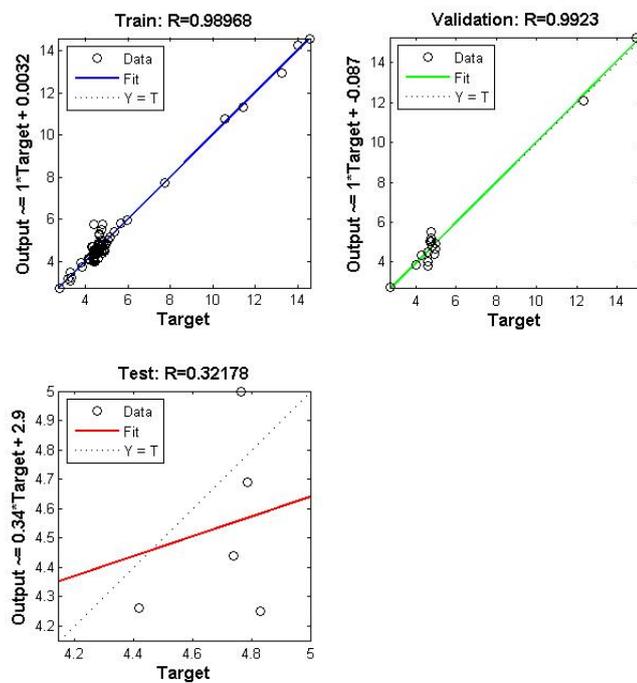


Figura 4.26: Gráfico do coeficiente de correlação da viscosidade para Resilient-Propagation com duas camadas.

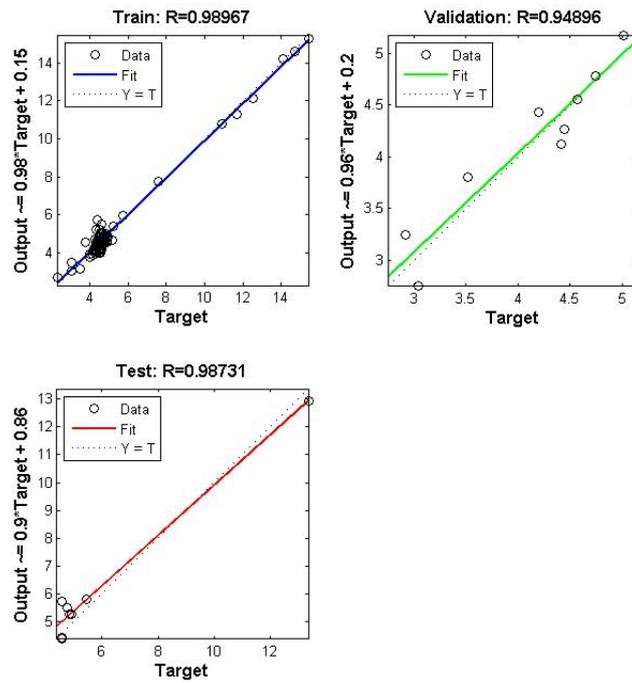


Figura 4.27: Gráfico do coeficiente de correlação da viscosidade para One-Step-Secant com duas camadas.

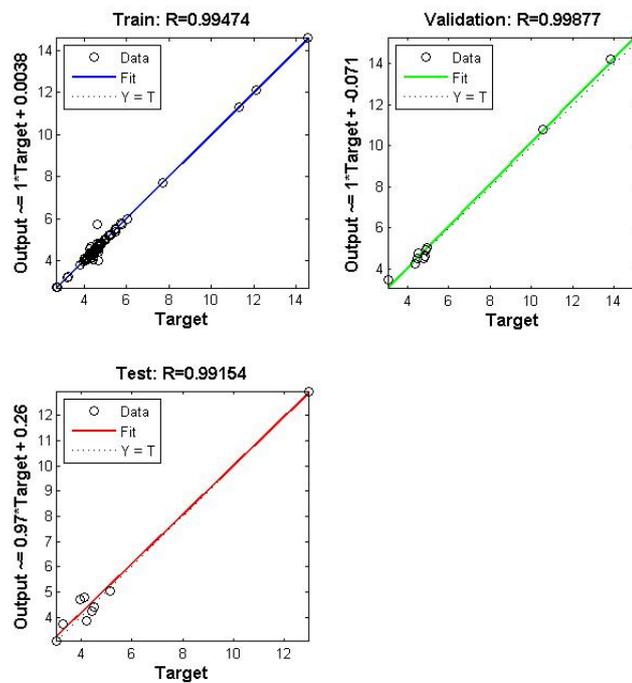


Figura 4.28: Gráfico do coeficiente de correlação da viscosidade para Levenberg-Marquardt com duas camadas.

## 5 Considerações Finais

Este capítulo apresenta os objetivos alcançados com a pesquisa e suas limitações. Além das contribuições científicas e sociais e os trabalhos futuros são apresentados.

### 5.1 Conclusões do trabalho

Este trabalho apresentou uma proposta de solução para predição de propriedades de qualidade do biodiesel, especificamente, índice de iodo e viscosidade cinemática, de modo a oferecer um método alternativo que renda menos custos para análise dessas propriedades.

No estudo foram usadas 98 amostras de ésteres de ácidos graxos para prever índice de iodo e viscosidade. Várias arquiteturas de rede foram implementadas e testadas com variações do algoritmo *backpropagation*, com o objetivo de se obter a melhor arquitetura de rede para prever cada propriedade.

Foram utilizadas técnicas de otimização e escolha de topologias candidatas de redes neurais, por meio de técnicas de validação cruzada por amostragem aleatória e validação cruzada por  $k$ -partições, além da adição do critério de parada antecipada e escolha do valor apropriado para o treinamento das redes neurais.

O principal objetivo proposto do trabalho de pesquisa foi avaliar a predição de propriedades de qualidade de biodiesel utilizando algoritmos de treinamento de aprendizado supervisionado de RNAs. E assim, contribuir com um estudo para se obter uma estrutura de rede com um algoritmo de treinamento que consiga bons resultados com melhor desempenho na predição.

Da avaliação desses algoritmos, pode-se perceber que as RNAs são uma boa ferramenta para predição de parâmetros de qualidade do biodiesel, as Figuras demonstraram índices de correlação da melhores arquitetura de rede para predição do índice de iodo e viscosidade com índice de correlação de 0.99 tanto para treinamento, validação e teste.

Deste modo, considera-se que com a avaliação dos algoritmos que é possível verificar a tendência da qualidade do biodiesel a partir de suas propriedades físico-químicas, deste modo reduzir tempo e custos para indústria de biocombustíveis.

## 5.2 Limitações

Embora o trabalho de pesquisa tenha alcançado seu objetivo principal, alguns pontos apresentam limitações podendo ser melhorados ou ampliados. Estes pontos são:

- Uma das principais limitações encontradas neste trabalho foram em relação a pouca quantidade de amostras e disposição espacial delas no conjunto de dados, dificultando a realização de treinamento. Já que é preciso assegurar que as amostras que carregam valores máximos e mínimos de cada variável de entrada, estejam distribuídas dentro dos conjuntos de treinamento, de modo a garantir um treinamento não tendencioso.
- Como a especificação da rede para resolução de um problema é feita de forma empírica, ou seja, uma busca de tentativa e erro, uma das dificuldades se encontra justamente na forma de elaborar a melhor heurística de solução do problema. Assim foi preciso testar vários tipos de combinações de redes e quantidade de neurônios diferentes para obter a melhor arquitetura de predição.
- Outra dificuldade limitante do trabalho é o fato da escolha do critério de parada ser diferente para cada propriedade, ou seja, uma mesma configuração de rede não pode ser aplicada a outra propriedade sem uma análise prévia de suas características.

## 5.3 Principais contribuições

Acredita-se que a pesquisa obteve as seguintes contribuições científicas e sociais:

- Avaliação do uso da técnica de RNAs aplicada a indústria dos biocombustíveis, por meio das predição das propriedades do biodiesel mostrando qual será a ten-

dência da qualidade do biodiesel ou até mesmo dispensar o processo de síntese do biocombustível no caso de uma matriz inapropriada para esta finalidade, o que possivelmente renderá custos e tempo para as indústrias de biocombustíveis.

- Elaboração de uma heurística de simulação de RNAs aplicado a predição de propriedades de qualidade do biodiesel.
- Predição da propriedade índice de iodo com a obtenção da melhor arquitetura de rede com duas camadas, por meio do algoritmo Levenberg-Marquardt, com coeficiente de correlação de 0.99 para treinamento, validação e teste.
- Predição da propriedade viscosidade com a obtenção da melhor arquitetura de rede com uma camada, por meio do algoritmo Levenberg-Marquardt, com coeficiente de correlação de 0.99 para treinamento, validação e teste.
- Artigo aceito para publicação intitulado *Application and Assessment of Artificial Neural Networks for Biodiesel Iodine Value Prediction* sobre o trabalho de pesquisa descrito nesta dissertação na ICANN 2015 : *International Conference on Artificial Neural Networks em Amsterdam, The Netherlands May, 14-15, 2015*
- Artigo aceito para publicação intitulado *Application of artificial neural networks to predict viscosity, iodine value and induction period of biodiesel focused on the study of oxidative stability* que utilizou a heurística elaborada nesse trabalho, na revista *FUEL* - 2014.

## 5.4 Trabalhos futuros

Os trabalhos futuros e melhoramento que podem ser derivados desta pesquisa são apresentados a seguir:

- Aplicação da abordagem implementada de RNAs para predição de outras propriedades de qualidade do biodiesel que estejam relacionadas com a composição físico-química do biodiesel.
- Aplicação e teste de outras técnicas de pré-processamento de dados para melhorar o desempenho dos algoritmos.

- 
- Fazer uso de técnicas de otimização como algoritmos genéticos ou simulated annealing para diminuir o escopo de tentativa de neurônios, diminuindo assim a quantidade de conexões numa rede MLP, como proposto em [50].

## Referências Bibliográficas

- [1] ADEWALE, P., MBA, O., DUMONT, M.-J., NGADI, M., AND COCCIARDI, R. Determination of the iodine value and the free fatty acid content of waste animal fat blends using FT-NIR. *Vibrational Spectroscopy* 72, 0 (2014), 72–78.
- [2] ALLEN, C., WATTS, K., ACKMAN, R., AND PEGG, M. Predicting the viscosity of biodiesel fuels from their fatty acid ester composition. *Fuel* 78, 11 (Sept. 1999), 1319–1326.
- [3] BAPTISTA, P., FELIZARDO, P., MENEZES, J. C., AND NEIVA CORREIA, M. J. Multivariate near infrared spectroscopy models for predicting the iodine value, CFPP, kinematic viscosity at 40 degrees C and density at 15 degrees C of biodiesel. *Talanta* 77, 1 (Oct. 2008), 144–51.
- [4] BASHEER, I., AND HAJMEER, M. Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods* 43, 1 (Dec. 2000), 3–31.
- [5] BEALE, M. H., HAGAN, M. T., AND DEMUTH, H. B. *Neural Network Toolbox™*. 2013.
- [6] BORSATO, D., MOREIRA, I., NOBREGA, M. M., MOREIRA, M. B., DIAS, G. H., AND BONA, E. Aplicação de redes neurais artificiais na identificação de gasolinas adulteradas comercializadas na região de londrina–paraná. *Quim. Nova* 32, 9 (2009), 2328–2332.
- [7] BRAGA, A. P., LUDERMIR, T. B., AND CARVALHO, A. F. *Redes neurais artificiais: teoria e aplicações*. LTC Editora, 2007.
- [8] BRASIL. Biodiesel: o novo combustível do brasil. Disponível em: <http://www.mme.gov.br/programas/biodiesel/>, (2004). Acesso em: 15 de julho de 2013.

- [9] BRASIL. Agência nacional do petróleo (anp). lei no 11.097. Disponível em: <http://www.anp.gov.br>, jan 2005. Acesso em: 26 agosto de 2013.
- [10] BRASIL. Programa nacional de produção e uso de biodiesel: Inclusão social e desenvolvimento territorial. Disponível em: <http://www.mda.gov.br/biodiesel>, (2011). Acesso em: 20 de maio de 2014.
- [11] BRASIL. Agência nacional do petróleo (anp). resolução no 14. Disponível em: <http://www.anp.gov.br>, maio 2012. Acesso em: 26 agosto de 2013.
- [12] CANDEIA, R. A. *Soybean Biodiesel: Synthesis, Degradation and Binary Mixtures*. PhD thesis, Doctoral dissertation, D. Sc. Thesis. Federal University of Paraíba. Brasil. Pag. 138, 2008.
- [13] ÇAY, Y., KORKMAZ, I., ÇIÇEK, A., AND KARA, F. Prediction of engine performance and exhaust emissions for gasoline and methanol using artificial neural network. *Energy 50* (Feb. 2013), 177–186.
- [14] CHAVARRIA-HERNANDEZ, J. C., AND PACHECO-CATALÁN, D. E. Predicting the kinematic viscosity of {FAMES} and biodiesel: Empirical models. *Fuel 124*, 0 (2014), 212 – 220.
- [15] DA SILVA, I. N., SPATTI, D. H., AND FLAUZINO, R. A. *Redes Neurais Artificiais para Engenharia e Ciências Aplicadas: curso prático*. Artliber.
- [16] DANTAS, M. B. *Biodiesel blends: flow properties, thermal and oxidative stability during storage and monitoring*. PhD thesis, Doctoral dissertation, D. Sc. Thesis. Federal University of Paraíba. Brasil. Pag. 138, 2006.
- [17] DUALIBE, A. K. *Combustíveis no Brasil: desafios e perspectivas*, 1nd ed. Synergia, Rio de Janeiro, 2012.
- [18] FREITAS, S. V., SEGOVIA, J. J., CARMEN MARTÍN, M., ZAMBRANO, J., OLIVEIRA, M. B., LIMA, A. S., AND A.P. COUTINHO, J. A. Measurement and prediction of high-pressure viscosities of biodiesel fuels. *Fuel 122* (Apr. 2014), 223–228.
- [19] GARCÍA, M., ALBA, J. J., GONZALO, A., SÁNCHEZ, J. L., AND ARAUZO, J. Density of alkyl esters and its mixtures: A comparison and improvement of predictive models. *Fuel 103* (Jan. 2013), 232–238.

- [20] GRAÇA, C. M. M. *Determinação do índice de iodo a partir da composição dos óleos*. PhD thesis, Dissertação de mestrado. Universidade de Aveiro. Pag. 98, 2010.
- [21] HAYKIN, S. *Neural Networks: A Comprehensive Foundation*, 2nd ed. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1998.
- [22] HOEKMAN, S. K., BROCH, A., ROBBINS, C., CENICEROS, E., AND NATARAJAN, M. Review of biodiesel composition, properties, and specifications. *Renewable and Sustainable Energy Reviews* 16, 1 (Jan. 2012), 143–169.
- [23] INSAUSTI, M., ROMANO, C., PISTONESI, M. F., AND BAND, B. S. F. Simultaneous determination of quality parameters in biodiesel/diesel blends using synchronous fluorescence and multivariate analysis. *Microchemical Journal* 108 (may 2013), 32–37.
- [24] JAHIRUL, M., BROWN, R., SENADEERA, W., O'HARA, I., AND RISTOVSKI, Z. The Use of Artificial Neural Networks for Identifying Sustainable Biodiesel Feedstocks. *Energies* 6, 8 (July 2013), 3764–3806.
- [25] JAHIRUL, M. I., SENADEERA, W., BROOKS, P., BROWN, R. J., SITU, R., PHAM, P. X., AND MASRI, A. R. An Artificial Neural Network (ANN) Model for Predicting Biodiesel Kinetic Viscosity as a Function of Temperature and Chemical Compositions. 1–6.
- [26] JAIN, A. K., MAO, J., AND MOHIUDDIN, K. Artificial neural networks: A tutorial. *Computer* 29, 3 (1996), 31–44.
- [27] KIVEVELE, T. T., MBARAWA, M. M., BERECZKY, A., LAZA, T., AND MADARASZ, J. Impact of antioxidant additives on the oxidation stability of biodiesel produced from Croton Megalocarpus oil. *Fuel Processing Technology* 92, 6 (2011), 1244–1248.
- [28] KNOTHE, G., KRAHL, J., AND GERPEN, J. V. *Biodiesel Handbook*, 2nd ed. AOCS Press, nov 2010.
- [29] KOHAVI, R., ET AL. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *IJCAI* (1995), vol. 14, pp. 1137–1145.
- [30] LOBO, I. P., FERREIRA, S. L. C., AND CRUZ, R. S. D. Biodiesel: quality parameters and analytical methods. *Química Nova* 32 (2009), 1596–1608.

- [31] MENG, X., JIA, M., AND WANG, T. Neural network prediction of biodiesel kinematic viscosity at 313K. *Fuel* 121 (Apr. 2014), 133–140.
- [32] MOSER, B. R. Biodiesel production, properties, and feedstocks. *In Vitro Cellular & Developmental Biology - Plant* 45, 3 (Mar. 2009), 229–266.
- [33] NOGUEIRA, C. A. Densities and Viscosities of Binary Mixtures of Babassu Biodiesel Cotton Seed or Soybean Biodiesel at Different Temperatures. 5305–5310.
- [34] OROMÍ-FARRÚS, M., VILLORBINA, G., ERAS, J., GATIUS, F., TORRES, M., AND CANELA, R. Determination of the iodine value of biodiesel using <sup>1</sup>H NMR with 1,4-dioxane as an internal standard. *Fuel* 89, 11 (nov 2010), 3489–3492.
- [35] PADHI, S. K. *Preparation and Characterization of Biodiesel from Non-Edible Oils*. PhD thesis, Doctoral dissertation, PhD thesis. National Institute of Technology, Rourkela, 2010.
- [36] PILOTO-RODRÍGUEZ, R., SÁNCHEZ-BORROTO, Y., LAPUERTA, M., GOYOS-PÉREZ, L., AND VERHELST, S. Prediction of the cetane number of biodiesel using artificial neural networks and multiple linear regression. *Energy Conversion and Management* 65 (jan 2013), 255–261.
- [37] PRECHELT, L. Automatic early stopping using cross validation: quantifying the criteria. *Neural Networks* 11, 4 (1998), 761–767.
- [38] RAMÍREZ-VERDUZCO, L. F., RODRÍGUEZ-RODRÍGUEZ, J. E., AND JARAMILLO-JACOB, A. D. R. Predicting cetane number, kinematic viscosity, density and higher heating value of biodiesel from its fatty acid methyl ester composition. *Fuel* 91, 1 (Jan. 2012), 102–111.
- [39] RAMOS, M. J., FERNÁNDEZ, C. M., CASAS, A., RODRÍGUEZ, L., AND PÉREZ, A. Influence of fatty acid composition of raw materials on biodiesel properties. *Bioresource technology* 100, 1 (Jan. 2009), 261–8.
- [40] REZAEI, M., BASRI, M., RAHMAN, R. N. Z. R. A., SALLEH, A. B., CHAIBAKHSH, N., AND MASOUMI, H. R. F. A multivariate modeling for analysis of factors controlling the particle size and viscosity in palm kernel oil esters-based nanoemulsions. *Industrial Crops and Products* 52, 0 (2014), 506–511.

- [41] REZENDE, J. A. D., M. A. L. Dependence of biodiesel fuel properties on the structure of fatty acid alkyl esters. *ISA Transactions* (1999), 209–214.
- [42] SALDANA, D. A., STARCK, L., MOUGIN, P., ROUSSEAU, B., FERRANDO, N., AND CRETON, B. Prediction of Density and Viscosity of Biofuel Compounds Using Machine Learning Methods. *Energy & Fuels* 26, 4 (Apr. 2012), 2416–2426.
- [43] SANFORD, S. D., WHITE, J. M., SHAH, P. S., WEE, C., VALVERDE, M. A., MEIER, G. R., GROUP, R. E., AND REPORT, C. Feedstock and Biodiesel Characteristics Report. 1–136.
- [44] SANLI, H., CANAKCI, M., AND ALPTEKIN, E. Predicting the higher heating values of waste frying oils as potential biodiesel feedstock. *Fuel* 115 (Jan. 2014), 850–854.
- [45] SANTOS, J. D. J. *Biodiesel of Babassu: Thermal Verification, Oxidative and Binary Blends*. PhD thesis, D. Sc. Thesis. Federal University of Paraíba, Brazil, 2008.
- [46] SHARON, H., JAYAPRAKASH, R., KARTHIGAI SELVAN, M., SOBAN KUMAR, D., SUNDARESAN, A., AND KARUPPASAMY, K. Biodiesel production and prediction of engine performance using SIMULINK model of trained neural network. *Fuel* 99 (sep 2012), 197–203.
- [47] SILVA, L. N. D. C. *Análise e síntese de estratégias de aprendizado para redes neurais artificiais*. PhD thesis, Dissertação de mestrado. Universidade Estadual de Campinas. Brasil. Pag. 210, 1998.
- [48] SOUZA, M. Deputados aprovam o aumento de mistura de biodiesel e etanol em combustíveis. Disponível em: <http://www2.camara.leg.br/camaranoticias/noticias/politica/472508-deputados-aprovam-aumento-de-mistura-de-biodiesel-e-etanol-em-combustiveis.html>. Acesso em: 10 de agosto de 2014.
- [49] TAFNER, M. A. *Estagiamento Automático do Sono Utilizando um Canal de EEG e uma Rede Neural Artificial com Alta Representação Cortical*. PhD thesis, Tese de Doutorado em Engenharia de Produção. Universidade Federal de Santa Catarina, 1999.

- 
- [50] YAMAKAZI, A. *Uma metodologia para otimização de arquiteturas e pesos de redes neurais*. PhD thesis, Doctoral dissertation, PhD thesis. Universidade Federal de Pernambuco, Recife, 2004.