

**UNIVERSIDADE ESTADUAL DE PONTA-GROSSA  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOENERGIA**

**ALINE REGINA WALKOFF**

**APLICAÇÃO DE REDES NEURAIS NO ESTUDO DO PERFIL DO ÁLCOOL  
ETÍLICO HIDRATADO COMBUSTÍVEL COMERCIALIZADO EM DIFERENTES  
REGIÕES DO ESTADO DO PARANÁ**

**PONTA GROSSA  
2016**

**ALINE REGINA WALKOFF**

**APLICAÇÃO DE REDES NEURAIS NO ESTUDO DO PERFIL DO ÁLCOOL  
ETÍLICO HIDRATADO COMBUSTÍVEL COMERCIALIZADO EM DIFERENTES  
REGIÕES DO ESTADO DO PARANÁ**

Dissertação apresentada para a obtenção do  
título de mestre em Bioenergia na  
Universidade Estadual de Ponta Grossa  
Área de concentração: Biocombustíveis

Orientador: Prof<sup>a</sup>. Dra. Sandra Regina  
Masetto Antunes  
Coorientador: Prof. Dr. Dionísio Borsato

**PONTA GROSSA**

**2016**

W186 Walkoff, Aline Regina  
Aplicação de redes neurais no estudo do perfil do álcool etílico hidratado combustível comercializado em diferentes regiões do estado do Paraná. / Aline Regina Walkoff. Ponta Grossa, 2016.  
49 f.

Dissertação (Mestrado em Bioenergia - Área de Concentração: Biocombustíveis), Universidade Estadual de Ponta Grossa.

Orientadora: Profa. Dra. Sandra Regina Masetto Antunes.

Coorientador: Prof. Dr. Dionísio Borsato.

1. Cana-de-açúcar. 2. Perceptron. 3. Mapas auto-organizáveis. 4. Parâmetros de conformidade. I. Antunes, Sandra Regina Masetto. II. Borsato, Dionísio. III. Universidade Estadual de Ponta Grossa. Biocombustíveis. IV.T.

CDD: 662.8

**TERMO DE APROVAÇÃO**

**ALINE REGINA WALKOFF**

**"APLICAÇÃO DE REDES NEURAIS NO ESTUDO DO PERFIL DO ÁLCOOL  
ETÍLICO HIDRATADO COMBUSTÍVEL COMERCIALIZADO EM  
DIFERENTES REGIÕES DO ESTADO DO PARANÁ"**

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Programa de Pós-Graduação em Bioenergia, área de concentração em Biocombustíveis, pela seguinte banca examinadora:

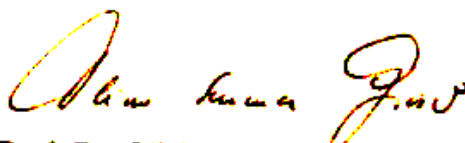


Orientadora: Profa. Dra. Sandra Regina Masetto Antunes  
UEPG/PR



Prof. Dr. Pedro Henrique Weirich Neto

UEPG/PR



Prof. Dr. Olívio Fernandes Galão  
UEL/PR

Ponta Grossa, 30 de março de 2016.

## **AGRADECIMENTOS**

Ao Prof Dionísio Borsato (UEL) por ter me convidado a conhecer o Programa de Pós-Graduação em Bioenergia, mostrando as pesquisas que realiza envolvendo redes neurais e a possibilidade de seu uso na área de biocombustíveis.

Aos professores da UEPG Sandra Regina Masetto Antunes e Pedro Henrique Weirich Neto pela admissão do projeto de pesquisa.

Aos alunos e alunas da UEL (especialmente Érica, Lívia, Hágata, César e Diego), que pesquisam junto ao Prof. Dionísio Borsato e que me ajudaram a operar os recursos computacionais necessários à execução deste trabalho, e também a interpretar e a descrever os resultados.

À Prof.<sup>a</sup> Sandra Regina Masetto Antunes, ao Prof. Dionísio Borsato e à Prof.<sup>a</sup> Carmen Luisa Barbosa Guedes (UEL), agradeço também as orientações e auxílios que me deram quando dos meus deslocamentos para outras cidades, bem como os ensinamentos e conselhos sobre a pesquisa científica geral e internacional.

Ao Prof. Paulo Rogério Pinto Rodrigues e equipe do laboratório de combustíveis da UNICENTRO (Guarapuava) por repassarem os dados necessários para este estudo.

## RESUMO

O etanol combustível deve estar adequado a parâmetros de controle de qualidade, tais como massa específica, teor de álcool, pH e condutividade elétrica. Suas características podem variar de com acordo sua região de comércio. Dados de análises físico-químicas foram coletados, de 998 amostras de etanol hidratado combustível comercializado em regiões designadas como “norte”, “centro-oeste” e “leste” do estado do Paraná. Os dados alimentaram redes neurais perceptron de múltiplas camadas e mapas auto-organizáveis, ambos os tipos de redes neurais artificiais, que classificaram as amostras conforme sua região de comercialização. A taxa de aprendizagem das redes neurais perceptron foi 0,10 e as amostras foram divididas aleatoriamente, sendo 70% para treinamento, 15% para teste e 15% para validação. Uma centena de redes foi treinada e o melhor desempenho foi obtido por uma rede neural com seis neurônios na camada oculta, a qual atingiu 85% de acerto para treinamento, 82% para teste e 84% para validação. A melhor configuração para mapas auto-organizáveis foi a de topologia 45 x 45 e 5000 épocas de treinamento, com uma taxa de aprendizagem final de  $6,7 \times 10^{-4}$ , uma relação de vizinhança final de  $3 \times 10^{-2}$  e um erro de quantização médio de  $2 \times 10^{-2}$ . Esta rede neural deu origem a um mapa topológico mostrando três grupos separados, cada um correspondendo a amostras de uma mesma região de comercialização. Quatro mapas de pesos, um para cada parâmetro, foram apresentados. Ambos os tipos de redes neurais tornaram possíveis a separação das amostras de acordo com a região de comercialização e concordaram que a massa específica foi um parâmetro relevante para a classificação.

Palavras-chave: cana-de-açúcar, perceptron, mapas auto-organizáveis, parâmetros de conformidade

## ABSTRACT

Ethanol fuel must comply with quality control parameters such as density, alcohol content, pH and electrical conductivity. Its characteristics may vary according to its trade regions. Physical-chemical analysis data were collected, from 998 hydrous ethanol fuel samples traded in regions designated as "north", "midwest" and "east" on the state of Paraná. The data fueled multilayer perceptron networks and self-organizing maps, both kinds of artificial neural networks, which classified the samples according to their trade regions. The perceptron networks learning rate was 0,10 and the samples were randomly divided, being 70% for training, 15% for testing and 15% for validation. One hundred networks were trained and the best performance was obtained by a network with six neurons in the hidden layer, which reached 85% of correction percentage for training, 82% for testing and 84% for validation. The self-organizing maps best configuration had a 45 x 45 topology and 5000 training epochs, with a final learning rate of  $6.7 \times 10^{-4}$ , a final neighborhood relationship of  $3 \times 10^{-2}$  and a mean quantization error of  $2 \times 10^{-2}$ . This neural network gave origin to a topological map depicting three separated groups, each one corresponding to samples of a same region of trade. Four maps of weights, one for each parameter, were presented. Both kinds of neural networks made possible the separation of samples according do their region of trade and agreed the density was a relevant parameter for the classification.

Keywords: sugarcane, perceptron, self organizing maps, compliance parameters

## LISTA DE FIGURAS

|             |  |    |
|-------------|--|----|
| Figura 1 –  | Evolução das não-conformidades de combustível no Brasil, em % das amostras coletadas, conforme o PMQC .....    | 15 |
| Figura 2 –  | Modelo estrutural básico de uma rede neural artificial .....   | 20 |
| Figura 3 –  | Dados da (a) massa específica, (b) pH, (c) teor alcoólico e (d) condutividade, para as amostras de etanol..... | 28 |
| Figura 4 –  | Matriz de gráficos contendo os parâmetros e o comportamento entre eles .....                                   | 29 |
| Figura 5 –  | Representação do erro obtido em relação ao número de ciclos de treinamento da rede MLP 4-6-3 .....             | 32 |
| Figura 6 –  | Representação da estrutura rede 02, MLP 4-6-3 .....  | 34 |
| Figura 7 –  | Erro de quantização da rede MAO em relação ao número de épocas .....   | 38 |
| Figura 8 –  | Gráfico de treinamento para a rede MAO de configuração 45x45   | 37 |
| Figura 9 –  | Mapa de pesos para a massa específica .....  | 38 |
| Figura 10 – | Mapa de pesos para o teor alcoólico .....  | 39 |
| Figura 11 – | Mapa de pesos para a condutividade elétrica.....   | 40 |
| Figura 12 – | Mapa de pesos para o pH.....   | 41 |



## LISTA DE TABELAS

|   |    |
|---|----|
| Tabela 1 – Estatística dos valores de parâmetros usados para caracterização de etanol .....                     | 30 |
| Tabela 2 – Taxa de acerto, e classe de funções utilizadas para as cinco melhores redes neurais MLP obtidas..... | 33 |
| Tabela 3 – Classificação de importância dos parâmetros conforme análise sensitiva .....                         | 35 |

## LISTA DE SIGLAS

|      |   |
|------|---|
| ABNT | Associação Brasileira de Normas Técnicas                    |
| AEHC | Álcool etílico hidratado combustível                        |
| ANP  | Agência Nacional do Petróleo, Gás Natural e Biocombustíveis |
| ASTM | American Society for Testing and Materials                  |
| BFGS | Broyden–Fletcher–Goldfarb–Shanno                            |
| ECS  | European Committee for Standardization                      |
| EHC  | Etanol hidratado combustível                                |
| MAO  | Mapas auto-organizáveis                                     |
| MLP  | Multilayerperceptron  |
| NBR  | Norma brasileira  |
| PMQC | Programa de Monitoramento da Qualidade dos Combustíveis     |

## SUMÁRIO

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>INTRODUÇÃO .....</b>  | <b>10</b> |
| <b>2</b> | <b>OBJETIVOS.....</b>  | <b>12</b> |
| 2.1      | OBJETIVO GERAL.....  | 12        |
| 2.2      | OBJETIVOS ESPECÍFICOS .....                                    | 12        |
| <b>3</b> | <b>REVISÃO BIBLIOGRÁFICA.....</b>                              | <b>13</b> |
| 3.1      | ETANOL HIDRATADO COMBUSTÍVEL.....                              | 13        |
| 3.1.1    | Parâmetros de qualidade do AEHC .....                          | 16        |
| 3.1.1.1  | <i>Massa específica e teor de álcool</i> .....                 | 17        |
| 3.1.1.2  | <i>Condutividade e pH</i> .....                                | 17        |
| 3.2      | REDES NEURAIS .....  | 19        |
| 3.2.1    | Redes perceptron de múltiplas camadas.....                     | 21        |
| 3.2.2    | Redes do tipo mapas auto-organizáveis.....                     | 22        |
| <b>4</b> | <b>MATERIAL E MÉTODOS.....</b>                                 | <b>24</b> |
| 4.1      | FORMAÇÃO DO BANCO DE DADOS .....                               | 24        |
| 4.1.1    | Teor de álcool e massa específica a 20°C .....                 | 25        |
| 4.1.2    | pH.....  | 25        |
| 4.1.3    | Condutividade elétrica .....                                   | 25        |
| 4.2      | REDES NEURAIS DO TIPO PERCEPTRON DE MÚLTIPLAS<br>CAMADAS ..... | 26        |
| 4.3      | REDES NEURAIS DO TIPO MAPAS AUTO-ORGANIZÁVEIS .....            | 26        |
| <b>5</b> | <b>RESULTADOS E DISCUSSÃO .....</b>                            | <b>27</b> |
| 5.1      | DADOS FÍSICO-QUÍMICOS DASAMOSTRAS.....                         | 27        |
| 5.2      | REDES NEURAIS DO TIPO PERCEPTRON DE MÚLTIPLAS<br>CAMADAS ..... | 30        |
| 5.3      | REDES NEURAIS DO TIPO MAPAS AUTO-ORGANIZÁVEIS .....            | 35        |
| <b>6</b> | <b>CONCLUSÃO .....</b>   | <b>42</b> |
| <b>7</b> | <b>SUGESTÕES PARA TRABALHOS FUTUROS .....</b>                  | <b>43</b> |
|          | <b>REFERÊNCIAS .....</b>                                       | <b>44</b> |

## 1 INTRODUÇÃO

O etanol é o biocombustível mais utilizado no Brasil, possuindo também relevância internacional, o que se percebe pela crescente evolução da tecnologia e pesquisa dirigida a aumentar a viabilidade de sua produção ou a se descobrir novas matérias-primas e meio de produção.

O controle da qualidade do etanol é parte integrante de sua cadeia produtiva e depende das características físico-químicas, que são avaliadas, no Brasil, de acordo com as normas estabelecidas pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP).

A ANP estabelece limites para alguns parâmetros físico-químicos, denominando-os parâmetros de conformidade, dentre eles encontram-se o pH, o teor alcoólico, a massa específica e a condutividade elétrica. Entretanto, esses parâmetros podem apresentar diferentes valores conforme matéria-prima, modo de produção, armazenamento, transporte, entre outros fatores. Para avaliar a contribuição de tais fatores nas diferentes características físico-químicas encontradas no etanol é necessário submeter os dados a métodos estatísticos. Um dos métodos são as redes neurais, que trabalham por reconhecimento de padrões estudando um banco de dados relativo às variáveis selecionadas. Os dados das variáveis são ligados aos fatores que se pretenda avaliar, e as redes neurais atuam com a finalidade de agrupar ou classificar amostras em categorias, conforme suas características diversas. As amostras semelhantes são colocadas numa mesma classificação ou em agrupamentos próximos.

O reconhecimento de padrões pode ser uma tarefa difícil de ser executada por meios estatísticos mais usuais, daí a relevância em se buscar experimentações com redes neurais artificiais, que buscam recriar em meio computacional a forma como o cérebro biológico opera em sua atividade de reconhecimento de padrões, capacidade esta inerente a vários seres vivos.

As redes neurais artificiais têm sido incorporadas a aplicações tão distintas como o reconhecimento automático de caracteres, identificação de padrões ou até mesmo à lógica de controle de uma máquina. A utilização dessas técnicas, assistidas por computador permite obter ganhos significativos de desempenho, auxiliando os especialistas em tomadas de decisões.

Vários autores têm utilizado resultados experimentais para o reconhecimento de padrões e classificação utilizando sistemas automatizados, o que é relevante na área de ciências exatas por possibilitar a obtenção de novas informações a partir de um mesmo conjunto de dados (LYRA, et al., 2010; RIBEIRO, et al., 2013). Dentro desse contexto, neste trabalho foram utilizadas as redes neurais artificiais para identificar padrões em amostras de etanol hidratado combustível, ligadas a características em comum, tais como sua região de comercialização.

## 2 OBJETIVOS

### 2.1 OBJETIVO GERAL

- Aplicar e adaptar redes neurais artificiais para classificar amostras de etanol hidratado combustível comercializadas nas regiões “norte”, “centro-oeste” e “leste” do estado do Paraná.

### 2.2 OBJETIVOS ESPECÍFICOS

- Constituir um banco de dados do perfil físico-químico do álcool etílico hidratado de acordo com a região de comercialização.
- Aplicar as redes neurais do tipo perceptron de múltiplas camadas e do tipo mapas auto-organizáveis na separação das amostras de etanol conforme a região de comercialização.
- Comparar os resultados fornecidos por ambas as redes.

### 3 REVISÃO BIBLIOGRÁFICA

#### 3.1 ETANOL HIDRATADO COMBUSTÍVEL

Os combustíveis derivados de hidrocarbonetos de petróleo, cuja utilização ainda é predominante, não possuem a vantagem do etanol, que pode ser feito a partir de recursos distribuídos de forma mais homogênea mundialmente (OWEN, 1995). O etanol é até o momento, o único combustível com capacidade de atender à crescente demanda mundial por energia renovável de baixo custo e de mitigar impactos ambientais negativos. As emissões gasosas com a queima do etanol são menores se comparadas às emissões da queima da gasolina, sendo ainda que parte do CO<sub>2</sub> emitido é reabsorvido pela própria cana (FELIPE, 2010; SILVA et al., 2008).

O etanol tem despertado de modo crescente a atenção de pesquisadores, empresas e governos, fato este decorrente de perspectivas de esgotamento das fontes não renováveis de combustíveis fósseis, competição relativa a preços de combustíveis, assim como preocupações de natureza ambiental relacionadas à emissão de substâncias que comprometem o meio ambiente (BASTOS, 2007).

O etanol permanece sendo o biocombustível mais utilizado no Brasil (SILVA et al., 2013). O país possui grande extensão territorial e diversificação climática, posicionando-se atualmente, no cenário internacional, como o maior produtor e exportador mundial de açúcar de cana e o maior produtor e consumidor de etanol. É também o único país a utilizar, em larga escala, o álcool como combustível alternativo ao petróleo. Na década de 70, o Brasil lançou o Programa Nacional do Álcool – PROALCOOL e desenvolveu tecnologia para abastecer com etanol, de forma extensiva, veículos até então movidos a gasolina. Atualmente é um dos países mais avançados, do ponto de vista tecnológico, na produção e no uso do etanol como combustível. O processo produtivo utilizado no Brasil é quase que exclusivamente a fermentação do mosto, constituído por caldo de cana e melaço (SILVA et al., 2006; SILVA et al., 2008).

Existem outras matérias-primas utilizadas no mundo para a obtenção do etanol. Uma fonte alternativa utilizada é o açúcar de beterraba (IGLESIAS & SESMERO, 2015). É possível também extrair amido de vegetais, como amido de milho (JOANNIS-CASSAN et al. 2014) e transformá-lo em glicose para sua posterior fermentação. Neste caso, enzimas fazem a liquefação e sacarificação do amido para

produzir glicose, a qual é fermentada por *Saccharomyces cerevisiae* para produzir o etanol (BURTON et al., 2013).

O etanol de segunda-geração vem surgindo a partir de tecnologia que permite converter celulose em glicose, para após realizar sua fermentação (WANG et al., 2016). Tal processo vem sendo adotado em escala comercial de forma crescente.

Além dos estudos relativos a matérias-primas e meios de obtenção, a caracterização química e controle de qualidade são também questões inerentes ao uso do etanol. O Brasil, sendo país pioneiro na sua exploração, possui normas relevantes quanto à regulamentação de sua comercialização e controle de qualidade.

No país, é a Agência Nacional do Petróleo, Gás Natural e Combustível (ANP) que exerce esta fiscalização, estabelecendo critérios de controle de qualidade do etanol e de outros combustíveis no Brasil, por meio de resoluções específicas. A Resolução nº 19/2015 define o etanol combustível como um biocombustível proveniente do processo fermentativo de biomassa renovável, destinado ao uso por motores a combustão interna, que possui como principal componente o etanol, o qual é especificado sob as formas de etanol anidro combustível e etanol hidratado combustível (EHC). O etanol hidratado combustível é definido como o etanol combustível que é destinado à utilização direta em motores a combustão interna, diferentemente do etanol anidro que é utilizado em misturas com a gasolina. A norma define também o etanol hidratado combustível *premium* (EHCP) como o que apresenta massa específica a 20°C variando de 799,7 a 802,8 kgm<sup>-3</sup> (BRASIL, 2015). A Resolução nº 19/2015 também especifica quais métodos de ensaios laboratoriais e parâmetros químicos serão considerados para o controle de qualidade do etanol. Os métodos indicados podem ser provenientes da Associação Brasileira de Normas Técnicas (ABNT), identificadas com o prefixo NBR, bem como de normas da *American Society for Testing and Materials* (ASTM), do *European Committee for Standardization* (ECS) e da *International Organization for Standardization* (ISO). Os métodos, geralmente, são os mesmos, apenas traduzidos e formatados de forma diversa conforme o órgão normativo (BRASIL, 2015). As normas NBR, em seu conteúdo, fazem referência a portarias publicadas pelo Inmetro, que é o Instituto Nacional de Metrologia, Qualidade e Tecnologia. Contudo, vários desses métodos de análise foram elaborados com base em outras matrizes,



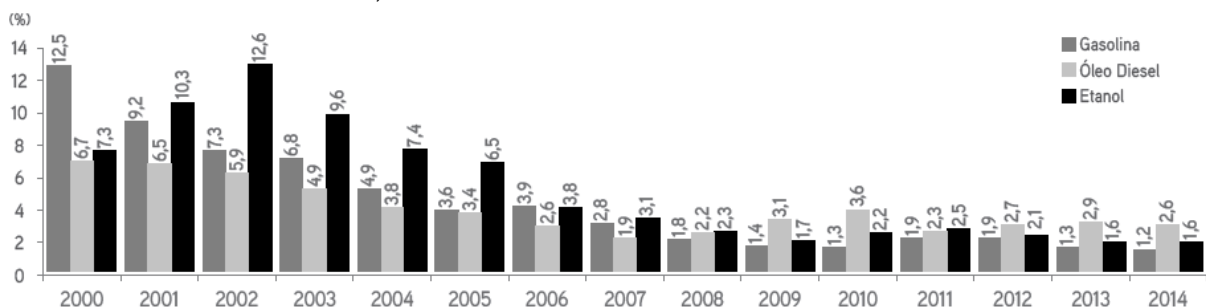
tais como as petrolíferas e de água, e não especificamente desenvolvidas a partir de testes com etanol.

O etanol precisa apresentar características físico-químicas adequadas para ser comercializado como combustível automotivo. Ao se avaliar os resultados de parâmetros químicos de controle da qualidade é possível inferir sobre suas prévias condições de armazenamento, contaminação por água e ácidos e presença de íons de metais, principalmente ferro e cobre (AGARWAL, 2007; GONZAGA et al., 2012).

A legislação brasileira prevê o controle de apenas algumas espécies químicas que podem estar presentes no etanol combustível, embora existam métodos para quantificação de outras. Substâncias como as que produzem efeitos tóxicos ao meio ambiente e à vida, tais como, os metais pesados poderiam ser quantificados e controlados (OLIVEIRA et al., 2009). A ANP mantém o Programa de Monitoramento da Qualidade dos Combustíveis (PMQC) para monitorar a qualidade da gasolina, do etanol e do óleo diesel comercializados nos postos revendedores. A criação do programa deu-se no ano de 1998 e, em 2005, passou a alcançar todas as unidades da federação. A cada mês são coletadas em torno de 18 mil amostras de gasolina, etanol hidratado e óleo diesel em postos revendedores de todas as regiões brasileiras, sorteados aleatoriamente os locais de coleta. As amostras são analisadas por mais de vinte laboratórios e instituições de pesquisa contratados pela ANP sob coordenação do laboratório que a própria Agência possui em Brasília.

A atuação da ANP subsidia as ações dos Ministérios Públicos federais e estaduais, Procons e Secretarias Estaduais e da Fazenda, os quais também participam no combate às adulterações. Os resultados do PMQC são publicados em boletins mensais. Todo esse trabalho vem ocasionando a queda constante nos índices de não conformidade dos combustíveis, como indica a Figura 1.

Figura 1 – Evolução das não-conformidades de combustível no Brasil, em % das amostras coletadas, conforme o PMQC. Fonte: ANP, 2016.



O último boletim do PMQC trouxe resultados de não-conformidade do etanol, no entanto, não trouxe informações sobre as possíveis origens dos valores inadequados, se por contaminação por metanol ou outras causas (BRASIL, 2016).

### 3.1.1 Parâmetros de Qualidade do AEHC

A norma da ANP vigente ao tempo da coleta de dados deste estudo, e a norma posterior, exigem dos distribuidores os controles dos parâmetros: aspecto, cor, massa específica, teor alcoólico, pH e condutividade elétrica (BRASIL, 2011; 2015). A Resolução 2015 da ANP exige dos produtores de etanol também os parâmetros: resíduo por evaporação, teor de hidrocarbonetos, teor de cloretos, teor de etanol, teor de água e teor de metanol. Na norma constam métodos e limites também para os teores de sulfato, ferro, sódio, cobre e enxofre, que deverão ser atendidos por todos os agentes que comercializam etanol, porém não exigindo boletim de conformidade relativo a esses parâmetros (BRASIL, 2015).

O Brasil conta com mais de 400 produtores de etanol a partir da cana-de-açúcar (GONÇALVES et al., 2013), sendo que as características do etanol podem se diferenciar conforme seus produtores ou distribuidores.

Métodos de análises físico-químicas para o controle da qualidade do etanol, ainda não contemplados pelos órgãos nacionais e internacionais, tais como ABNT, ASTM, ISO e CEN, vêm sendo apresentados e discutidos na literatura, pois é relevante o conhecimento das diferentes espécies químicas que possam estar presentes. Já foram feitos estudos com técnicas de espectrometria, cromatografia ou eletroquímica para detectar diversos compostos orgânicos e inorgânicos, dentre os quais se destacam: chumbo, alumínio, cádmio, ésteres e aldeídos aromáticos. A alta confiabilidade e detecção de substâncias em baixos limites por esses métodos instrumentais vêm possibilitando alteração e adição de novas normas oficiais quanto ao controle de qualidade do etanol combustível (OLIVEIRA et al., 2009).

### 3.1.1.1 Massa específica e teor de álcool

Em amostras de álcool hidratado combustível contendo apenas etanol e água, o valor do teor alcoólico é inversamente proporcional ao da massa específica, pois quanto maior a presença de etanol, menor a densidade total da amostra. Ademais, o método mais usual e recomendado pela ANP (NBR 5992) para a realização do ensaio, obtém o valor do teor alcoólico por meio de cálculo que tem como base o valor da massa específica. Assim, pode-se dizer que ambos os parâmetros são colineares numa representação gráfica.

A massa específica está relacionada ao teor de água contido no etanol hidratado combustível. O teor de álcool avalia se há contaminações com hidrocarbonetos, metanol e outros alcoóis (TANAKA et al., 2011).

Nos últimos anos a adulteração de etanol com adição de metanol tem sido causa de preocupação para a ANP, pois apesar de o metanol possuir propriedades físico-químicas semelhantes ao etanol, aquele composto orgânico pode causar problemas de saúde (SILVA et al., 2013). O metanol possui densidade menor que o etanol, portanto, uma amostra contaminada apresentará menor resultado para massa específica. Porém, a adulteração com metanol não é percebida através do ensaio de massa específica se houver água adicionada à amostra, pois esta possui maior densidade que o etanol compensando o metanol que também foi adicionado.

Os valores de conformidade em 2013 para a massa específica estavam entre 807,6 e 811,0 kg m<sup>-3</sup> e para o teor alcoólico entre 92,5 e 93,8% em massa (°INPM). A partir do ano de 2015 a ANP passou a admitir limites um pouco mais amplos, de 805,2 a 811,2 kg m<sup>-3</sup> para a massa específica e 92,5 e 94,6% em massa (°INPM) para o teor alcoólico.

### 3.1.1.2 Condutividade e pH

Ambos os parâmetros relacionam-se à presença de ácidos e potencial corrosivo do etanol. Um contaminante orgânico comum no etanol é o acetato de etila, proveniente da reação do etanol com o ácido acético. Este advém do processo produtivo do etanol, nas fases de fermentação dos açúcares, da destilação ou devido a contaminantes. A acidez no etanol é um parâmetro químico que mensura o

conteúdo total de ácidos tituláveis (GONZAGA et al., 2012). Outra substância que pode estar presente e aumentar a acidez no etanol é o dióxido de carbono (GONZAGA et al., 2012) que, como esclarece Dalmolin et al. (2006), é mais facilmente solubilizado no etanol do que na água.

A determinação de ácido acético é usualmente feita por cromatografia, no entanto, outros métodos vêm sendo desenvolvidos, tais como titulação coulométrica (GONZAGA et al. 2012) e voltamétrico (CAETANO et al., 2013). A acidez influencia na redução do pH (CAETANO et al., 2013; GONZAGA et al., 2012), por isso a norma da ANP pode requerer o controle do pH para distribuidores de combustíveis mas deixar de exigir o índice de acidez.

Segundo Ballerini (2006), a presença de ácido acético em etanol leva a um aumento da condutividade do mesmo, já que uma solução de  $0,1 \text{ mol L}^{-1}$  de ácido acético possui condutividade de  $41 \mu\text{S m}^{-1}$  e o etanol cerca de  $1 \times 10^{-3} \mu\text{S m}^{-1}$ .

O aumento da condutividade facilita o fenômeno da corrosão de partes metálicas com as quais o etanol esteja em contato, caso haja sais metálicos no etanol que favoreçam a troca iônica (DELGADO et al., 2007). Trata-se da corrosão galvânica, que ocorre pelo contato entre dois metais em um meio com eletrólitos e condutividade elétrica (BALLERINI, 2006). Portanto, um baixo pH favorece a corrosão do etanol sobre o tanque de combustível e outras peças do automóvel (GONÇALVES et al., 2011). A corrosão também danifica motores, peças e equipamentos de usinas onde o etanol seja produzido ou mesmo utilizado (OLIVEIRA, 2009; HUSSAN et al., 2013).

O etanol é muito higroscópico, pode absorver íons solubilizados em água que irão contaminá-lo aumentando sua condutividade (HUSSAN et al., 2013). Se ficar armazenado em tanque em contato com o ar atmosférico, pode absorver água. Tal efeito acontece, por exemplo, em colheitadeiras agrícolas, que utilizam misturas de combustíveis de diesel e etanol. Pode ocorrer que o combustível permaneça parado na bomba injetora por meses; se o etanol estiver com características impróprias, as partes metálicas serão danificadas (LA HARPE, 1988).

Para a condutividade elétrica desde o ano de 2015 o limite máximo aceito pela ANP é de  $300 \mu\text{S m}^{-1}$ , sendo que à época de coleta dos dados destes estudos era de  $389 \mu\text{S m}^{-1}$  (BRASIL, 2013). O pH em ambas as normas é limitado entre 6,0 e 8,0.

### 3.2 REDES NEURAIS

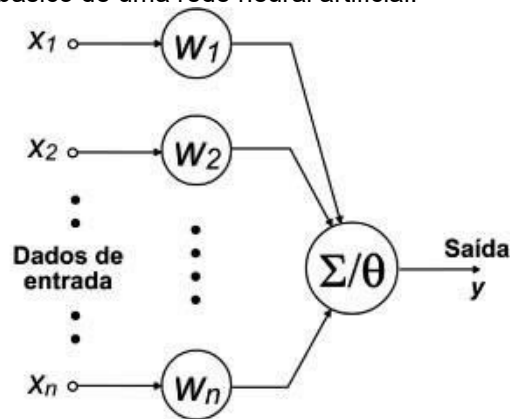
A classificação quimiométrica de amostras, na maioria dos casos, é bastante complexa, principalmente se os padrões forem descritos por um grande número de variáveis independentes; quando isso ocorre é necessário utilizar sistemas automatizados. Um dos sistemas são as redes neurais artificiais, que procuram modelar, mesmo que primitivamente, as operações lógicas pelas quais o cérebro executa as mais variadas tarefas (DEISINGH, 2004; LIAO, 2004).

Ao longo de milhares de anos, devido à evolução e à adaptação do sistema nervoso central humano, o reconhecimento e a classificação de padrões é uma tarefa naturalmente realizada (HAIR, 2009; HAYKIN, 2001). As redes neurais artificiais surgiram como uma alternativa a essa tarefa, operando por meios computacionais. Em circuitos de silício os impulsos elétricos ocorrem na ordem de nanossegundos ( $10^{-9}$  s), enquanto que no cérebro humano acontecem na ordem de milissegundos ( $10^{-3}$  s). Os sistemas artificiais geralmente são configurados para utilização de menor número de neurônios, já o cérebro humano compensa a menor taxa de operação com um número elevado de neurônios e conexões maciças entre si (HAYKIN, 2001).

As redes neurais artificiais e os neurônios cerebrais assemelham-se por adquirirem conhecimento através de um processo de aprendizagem e por armazenarem conhecimento conforme um peso sináptico, ou seja, a força de conexão entre neurônios (HAYKIN, 2001). Cada neurônio artificial pode receber vários estímulos de entrada e gerar vários estímulos de saída, sendo que este fluxo de estímulos é propagado através de uma rede de conexões, tal qual no modelo biológico. A intensidade de propagação dos estímulos entre uma determinada conexão é feita através de pesos de ponderação. Os sinais ponderados recebidos são somados e modificados através de uma função de transferência (KOHONEN, 1997).

Um modelo estrutural básico para uma rede neural é representado na Figura 2. Os dados de entrada são representados pelas variáveis "x", sendo que todas transmitem suas informações para uma função de transferência representada pela soma dessas informações e que poderá emitir um sinal de saída "y". As variáveis de entrada apresentam conexão com a função de transferência de acordo com os valores de peso, representados por "w".

Figura 2 – Modelo estrutural básico de uma rede neural artificial.



Fonte: Fiorinet al.(2011)

Existem vários tipos de redes neurais tais como perceptron de múltiplas camadas, redes de base radial, mapas auto-organizáveis, dentre outras (HAYKIN, 2001; BISHOP, 2007).

No campo das ciências humanas e sociais aplicadas, os mapas auto-organizáveis podem atuar, por exemplo, na mensuração bem-estar numa sociedade (LUCCHINI et al., 2013). Outro estudo buscou agrupar países estabelecendo sua semelhança conforme indicadores sociais (RENDE et al., 2011). Na economia, relacionam finanças e sociedade (LOUIS et al., 2013). Em estudos comportamentais as redes neurais podem aprender e prever ações humanas em situações tais como a de dirigir um carro (VANDERHAEGEN et al., 2014).

Na medicina as redes neurais foram utilizadas no estudo da absorção intestinal de insulina (KAMEI et al., 2012) e na representação de padrões de atividade cerebral, (EBADZADEH et al., 2003; HAXBY et al., 2014). Na oceanografia pode-se verificar o crescimento de biomassa de fitoplâncton conforme características da região marítima (SMITH et al., 2014). Na geologia verificaram estruturas de argila em rochas (MELLER et al., 2014). Na astronomia as redes neurais permitem a mineração de grande quantidade de dados (GEACH, 2012).

Para as áreas de química e agronomia, as redes neurais artificiais vêm sendo utilizadas na pesquisa de quimioinformática (BONACHERA et al., 2012), na discriminação de amostras de café conforme a região de plantio (SIERRA et al., 2005; BORSATO et al., 2011), no estudo do perfil de gasolinas comercializadas em diferentes regiões (SILVA et al., 2015), na segmentação de amostras de soja pela determinação do conteúdo de compostos inorgânicos (CREMASCO et al., 2016) e

na classificação de amostras de etanol conforme a destilaria de origem (SPACINO, 2011).

### 3.2.1 Redes perceptron de múltiplas camadas

A rede perceptron de múltiplas camadas, ou *multilayer perceptron* (MLP) trata-se de rede neural do tipo supervisionada, que é treinada utilizando amostras conhecidas e que faz o ajuste de seus parâmetros internos, denominados pesos, de forma que o erro vá diminuindo à medida que a rede compara os resultados que está prevendo aos resultados que constam da base de dados (HAYKIN, 2001; BISHOP, 2007).

Em seu funcionamento a rede recebe as informações das variáveis de entrada e os sinais são transmitidos através das camadas ocultas, até se obter uma resposta de saída. Os sinais que se transmitem com maior intensidade correspondem a vetores de pesos maiores. A rede verifica se a saída, ou seja, a região da amostra de etanol corresponde ao que lhe foi informado. Em caso de erro a rede retorna sinal para que sejam feitos ajustes nos pesos de transmissão de sinais, trata-se da retropropagação do sinal (HAYKIN, 2001; BISHOP, 2007).

Os neurônios da camada oculta nas redes perceptron agem como detectores de características e aprendem por retropropagação porque, à medida que o processo de aprendizado avança, gradualmente descobrem as peculiaridades que caracterizam os dados (HAYKIN, 2001; BISHOP, 2007). Funções que recebem as informações dos sinais formam as camadas ocultas e são as responsáveis pelo ajuste dos pesos. O erro decai conforme os pesos são ajustados. As redes perceptron são supervisionadas porque funcionam por meio de um algoritmo de retropropagação, que possibilita a aprendizagem ao realizar as correções dos erros de saída (HAYKIN, 2001; BISHOP, 2007). A rede pode ser programada para utilizar parte das amostras para seu treinamento e o restante para teste e validação dos resultados, fornecendo ao fim a porcentagem de acerto para cada etapa (BORSATO et al., 2009; BORSATO et al., 2011).

Romagnoli et al. (2016) utilizaram as redes neurais do tipo perceptron a fim de identificar a região de origem de 204 amostras de etanol, pertencentes às regiões “norte” e “leste” do estado do Paraná. Foi informado para as redes neurais testadas um banco de dados contendo os valores de massa específica, teor de

álcool, pH e condutividade elétrica das amostras a serem classificadas. Uma rede neural perceptron contendo uma camada oculta com oito neurônios apresentou o melhor desempenho, com 95% de acerto para treinamento, 96% para teste e 96% para validação.

### 3.2.2 Redes do tipo mapas auto-organizáveis

Os mapas auto-organizáveis (MAO), ou de Kohonen, são redes neurais que utilizam somente as variáveis de entrada para encontrar similaridades entre as amostras. Não há retropropagação dos sinais de saída para ajustar os pesos conforme haja erros, portanto se tratam de redes não supervisionadas, diferentemente das redes perceptron. Seu aprendizado é competitivo já que se trata de uma competição dos neurônios pelas variáveis de saída (KOHONEN, 1997).

Os mapas auto-organizáveis são capazes de mapear um conjunto de dados de um espaço de entrada, contido em um conjunto finito de neurônios organizados em arranjo unidimensional ou bidimensional, sendo adequados para a tarefa de seleção de características (KOHONEN, 1997). Funcionam comprimindo variáveis num plano bidimensional, resultando num mapa topológico onde se poderá verificar a presença de agrupamentos, também chamados de *clusters*, das amostras a que se referem (LUCCHINI et al., 2013). Os agrupamentos são identificados visualmente sendo que amostras próximas indicam uma relação de vizinhança e similaridade de características (KOHONEN, 1997).

Num estudo realizado por Spacino et al. (2013), 218 amostras de etanol comercializadas na cidade de Londrina foram analisadas quanto a seu teor de etanol, massa específica, pH e condutividade elétrica, de acordo com métodos recomendada pela ANP. Esses parâmetros formaram o banco de dados para alimentar redes neurais de topologia 4x7 com 1000 ciclos de treinamento. Dois experimentos foram realizados, sendo que em cada um a rede dividia parte das amostras para executar um treinamento e o restante para ser utilizado como validação. No total, foram apresentadas duas redes de treinamento e duas de validação, todas com mesma configuração. As redes neurais permitiram o agrupamento das amostras num mapa de neurônios, separando-as em sete grupos de amostras de acordo com sua destilaria de origem, com 100% de precisão na classificação das amostras. Portanto, as redes neurais artificiais, ao examinarem o



banco de dados de análises de etanol, podem indicar que existe um padrão inerente àquelas amostras, ligado às características que possuem em comum, tais como a região de sua matéria-prima ou seu processo produtivo.

Os mapas auto-organizáveis também permitem a geração de mapas de pesos para cada variável de entrada, o que permite a compreensão em detalhes sobre como cada fator estudado interfere na segmentação observada (LINK, 2011).

## 4 MATERIAL E MÉTODOS

### 4.1 FORMAÇÃO DE BANCO DE DADOS

Durante o período compreendido entre os meses de abril e dezembro de 2014 foram coletadas e analisadas 998 amostras de álcool etílico hidratado combustível por três laboratórios de controle de qualidade de combustíveis do estado do Paraná.

As amostras foram provenientes das distribuidoras de combustíveis, as quais providenciaram as análises de parâmetros de conformidade nos laboratórios mais próximos, de acordo com sua localização. As distribuidoras tendem a fornecer seu combustível aos postos de abastecimento mais próximos, por isso pode-se identificar regiões de comercialização. Neste trabalho as amostras coletadas foram provenientes das regiões do estado do Paraná definidas como “norte”, “leste” e “centro-oeste”, com base nesse critério de localização de laboratórios e distribuidoras.

O Laboratório de Pesquisas e Análises de Combustíveis da Universidade Estadual de Londrina, localizado na cidade de Londrina, analisou 326 amostras, sendo que para fins deste estudo, tais amostras foram denominadas como pertencentes à região “norte” do estado do Paraná.

O Laboratório Chronion Análises Químicas e Comércio Ltda., localizado na cidade de Quatro Barras, analisou 420 amostras que corresponderam à região “leste”.

O Laboratório de Combustíveis da Universidade Estadual do Centro Oeste, situado na cidade de Guarapuava, analisou 252 amostras denominadas como da região “centro-oeste”.

As amostras foram submetidas a ensaios de pH, teor alcoólico, condutividade elétrica e massa específica a 20°C. Os resultados das análises formaram a base de dados que alimentou tanto as redes neurais do tipo perceptron de múltiplas camadas como as do tipo mapa auto-organizáveis.

#### 4.1.1 Teor de álcool e massa específica a 20 °C

Para determinação do teor alcoólico e massa específica a 20 °C a norma seguida foi a NBR 5992 (ABNT, 2008), equivalente à ASTM D4052-11 (ASTM, 2011). A norma define a massa específica como a massa por unidade de volume de uma substância a uma determinada temperatura.

O valor de massa específica anotado é relativo à temperatura em que foi realizado o ensaio, sendo corrigido para 20 °C. A correção foi feita através de um *software* denominado "Programa de Tabelas Alcoolométricas", que é fornecido juntamente com a norma NBR 5992.

O teor alcoólico foi calculado por equação indicada no método contido na norma NBR 5992, a partir do resultado da massa específica a 20 °C, sendo seu resultado expresso em porcentagem de massa ou volume de etanol em relação ao total da amostra. Neste estudo a porcentagem foi determinada em relação à massa, ou seja em °INPM. Não há valores de referência desses parâmetros para o etanol.

#### 4.1.2 pH

O pH foi determinado conforme a norma NBR 10891, que especifica o método potenciométrico. Este método pode ser utilizado em etanol hidratado combustível, contendo até 3%, em volume, de hidrocarbonetos. O valor de pH obtido por este método não é compatível a valores de pH em soluções aquosas. O resultado é obtido pela diferença de potencial entre os eletrodos de pH de medida e de referência prata/cloreto de prata.

#### 4.1.3 Condutividade elétrica

Foi determinada de acordo com a norma NBR 10547, equivalente à ASTM D1125-11 (ASTM, 2014). A condutividade elétrica é a capacidade de um material em conduzir corrente elétrica por íons presentes na solução, fazendo-se uso de célula de condutividade elétrica que opera com sensor para medição da temperatura da amostra. O resultado é expresso em números inteiros na unidade de micro Siemens por metro ( $\mu\text{S m}^{-1}$ ). Não há tendências ou valores de referência.

## 4.2 REDES NEURAI DO TIPO PERCEPTRON DE MÚLTIPLAS CAMADAS

Foi utilizado o módulo de redes neurais do software Statistica 9.0 e a entrada dos parâmetros foi realizada na seguinte ordem: massa específica a 20°C, teor alcoólico, pH e condutividade elétrica. Foi acionado o modo automático padrão para a seleção de funções e número de neurônios. As redes foram treinadas com 70% das amostras, e utilizaram 15% das amostras para teste e 15% para validação. A escolha das amostras em cada inicialização da rede foi feita de forma randômica e a taxa de aprendizagem inicial foi de  $5 \times 10^{-2}$ . As variáveis categóricas foram as regiões de comercialização e as variáveis contínuas foram os quatro parâmetros químicos.

As redes apresentaram uma camada oculta, a qual possuiu número de neurônios variando entre 3 e 10 e executaram no máximo 200 ciclos de treinamento. O módulo de redes neurais selecionou também randomicamente as funções de erro, ativação da camada oculta e de saída dentre as disponibilizadas. Para o erro estavam disponíveis as funções de soma dos quadrados e entropia; para a ativação da camada oculta as funções foram identidade, logística, tangente e exponencial; para a saída as funções foram identidade, logística, tangente, exponencial, e softmax. O algoritmo de retropropagação selecionado foi o BFGS. Todas as funções e algoritmos fazem parte do módulo de redes neurais do *software* utilizado, já estando nele previamente adicionadas.

## 4.3 REDES NEURAI DO TIPO MAPAS AUTO-ORGANIZÁVEIS

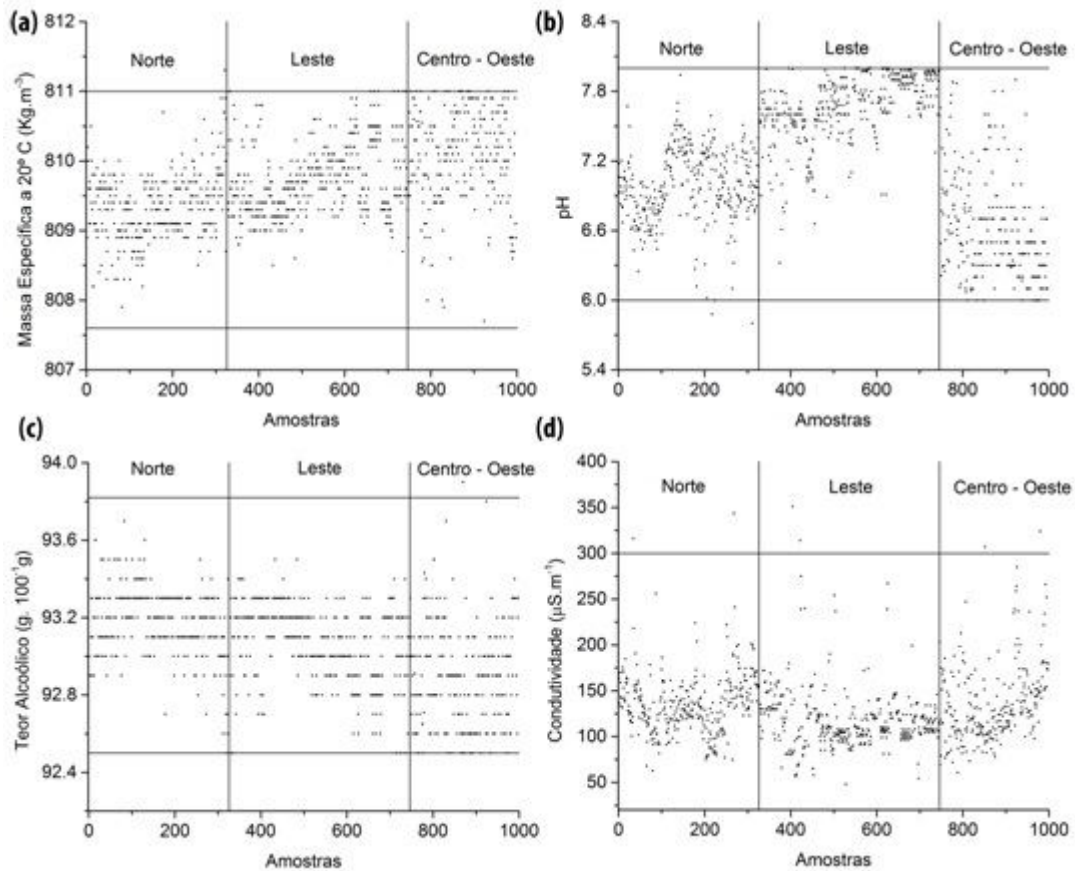
Foi utilizado o módulo de redes neurais do *software* Matlab R2007 e a entrada dos parâmetros foi realizada na seguinte ordem: massa específica a 20°C, teor alcoólico, pH e condutividade elétrica. Os dados foram processados em um computador Intel Core i7-4790 de 3,60 GHz e 32 GB de memória RAM. As amostras da região “leste” foram designadas pela letra L, as da “centro-oeste” pela letra C e as da “norte” pela letra N.

## 5 RESULTADOS E DISCUSSÃO

### 5.1 DADOS FÍSICO-QUÍMICOS DAS AMOSTRAS

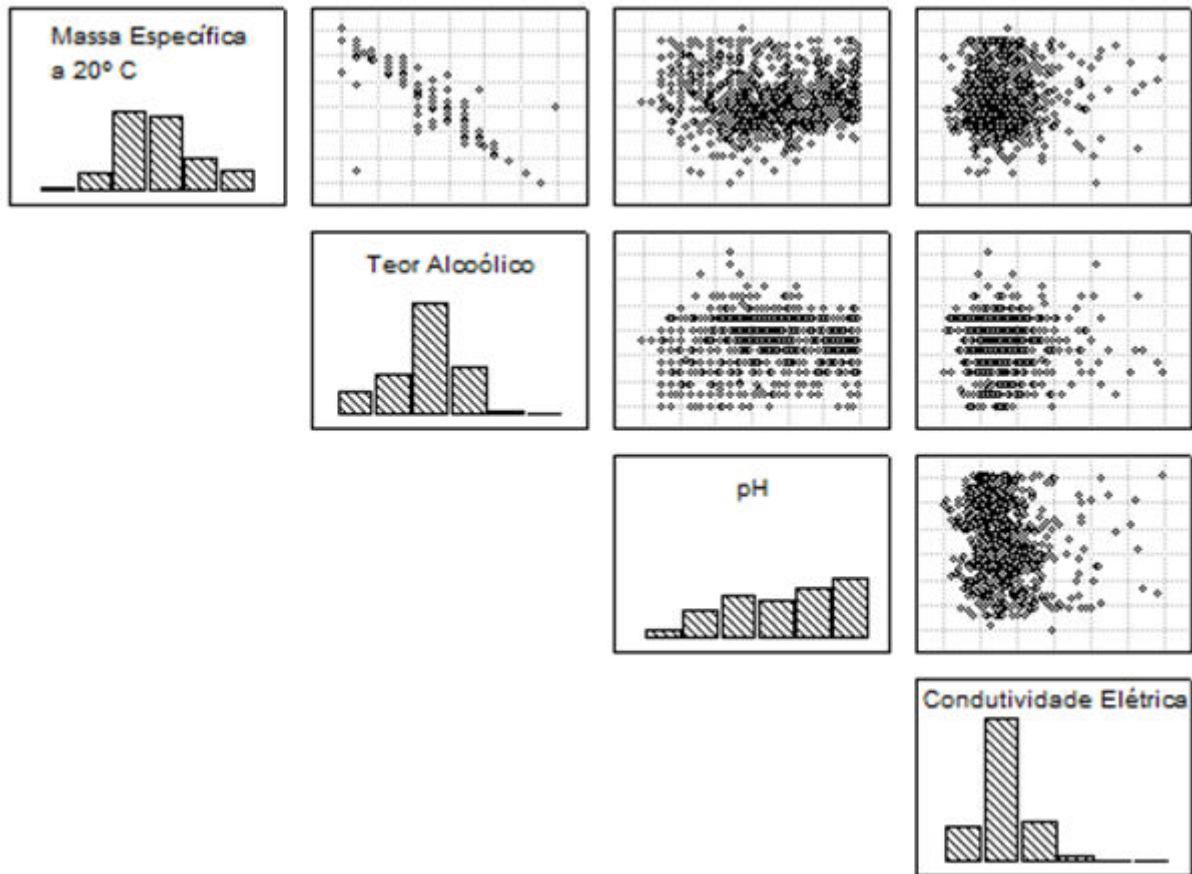
A Figura 3 mostra os dados das 998 amostras, em ordem cronológica por região, para a massa específica, pH, teor alcoólico e condutividade elétrica. Para o pH, os resultados da região “centro-oeste” estão com valores mais dispersos, comparando-se com os das regiões “leste” e “norte”. As linhas horizontais estabelecem os limites superior e inferior dos parâmetros de conformidade utilizados. Os valores de conformidade para a massa específica estão entre 807,6 e 811,0 kg m<sup>-3</sup>, para o pH entre 6,0 e 8,0, para o teor alcoólico entre 92,5 e 93,8% em massa (°INPM); esses eram os valores vigentes conforme a resolução da ANP à época dos ensaios (BRASIL, 2013), sendo que a resolução da ANP nº 19 do ano de 2015 passou a adotar limites um pouco mais amplos para a massa específica e o teor alcoólico. Para a condutividade foi adotado um limite máximo de 300 µS m<sup>-1</sup>, valor este fixado pela ANP e menor que os das normas antecedentes. Na região “norte”, uma amostra está acima do limite superior para a massa específica, duas acima do superior para a condutividade e duas abaixo do limite inferior para o pH. Para o teor alcoólico, uma amostra do “centro-oeste” apresentou resultado acima do limite superior. Para a condutividade, duas amostras de cada região ultrapassaram o limite superior (BRASIL, 2013; 2015).

Figura 3 – Dados da (a) massa específica, (b) pH, (c) teor alcoólico e (d) condutividade, para as amostras de etanol.



A Figura 4 mostra as distribuições gráficas dos valores obtidos para as correlações entre os parâmetros para as 988 amostras. Cada distribuição permite notar a relação e comportamento entre os dois parâmetros referentes a cada eixo. Quando se correlacionam os valores da massa específica com os de teor de álcool, o comportamento da distribuição das amostras é próximo da linearidade com  $R^2$  0,8932. Já as outras correlações mostraram um caráter tipicamente não linear com  $R^2$  máximo de 0,06. As relações não lineares entre os demais parâmetros mostram que as redes neurais podem ser aplicadas neste estudo de caso (RITTER, 1995).

Figura 4 – Matriz de gráficos contendo os parâmetros e o comportamento entre eles, indicando linearidade entre a massa específica e o teor alcoólico.



A Tabela 1 especifica os valores mínimos e máximos de cada parâmetro para as 998 amostras, bem como suas médias e desvios padrão, de acordo com sua região de comercialização. Os desvios padrão estão com maiores valores para a condutividade, pois seus resultados tendem a estar mais dispersos em relação ao valor médio. Nos valores das médias, aqueles com letras iguais indicadas ao lado, são os que não apresentam diferença significativa de acordo com o teste de Tukey. Porém, de acordo com o teste de Levenes, o valor do p foi menor que 0,05, indicando que a hipótese de homogeneidade de variância não deve ser aceita. Portanto, o teste de Tukey deve ser desconsiderado por não poder ser aplicado para o caso em estudo. As redes neurais vieram a trazer informações sobre os dados que não seriam facilmente obtidas por métodos estatísticos convencionais.

Tabela 1 – Dados estatísticos dos valores dos parâmetros usados para caracterização de etanol.

| Todas as regiões      | Massa específica (kg m <sup>-3</sup> ) | pH                   | Teor alcoólico(g 100 <sup>-1</sup> g)  | Condutividade (µS m <sup>-1</sup> ) |
|-----------------------|--|----------------------|--|-------------------------------------|
| Valor mínimo          | 807,7000                               | 5,8000               | 92,5000                                | 48,0000                             |
| Valor máximo          | 811,3000                               | 8,0000               | 93,9000                                | 351,0000                            |
| Valor médio           | 809,6880                               | 7,1801               | 93,0672                                | 128,7298                            |
| Desvio padrão         | 0,6194                                 | 0,5535               | 0,2291                                 | 36,7805                             |
| Região “norte”        | Massa específica (kg m <sup>-3</sup> ) | pH                   | Teor alcoólico (g 100 <sup>-1</sup> g) | Condutividade (µS m <sup>-1</sup> ) |
| Valor mínimo          | 807,9000                               | 5,8000               | 92,5000                                | 62,8000                             |
| Valor máximo          | 811,3000                               | 7,94000              | 93,7000                                | 343,4000                            |
| Valor médio           | 809,3439 <sup>a</sup>                  | 6,98675 <sup>a</sup> | 93,1804 <sup>a</sup>                   | 136,0692 <sup>a</sup>               |
| Desvio padrão         | 0,4905                                 | 0,3128               | 0,1736                                 | 33,5836                             |
| Região “leste”        | Massa específica (kg m <sup>-3</sup> ) | pH                   | Teor alcoólico(g 100 <sup>-1</sup> g)  | Condutividade (µS m <sup>-1</sup> ) |
| Valor mínimo          | 808,5000                               | 6,3200               | 92,5000                                | 48,0000                             |
| Valor máximo          | 811,0000                               | 8,0000               | 93,5000                                | 351,0000                            |
| Valor médio           | 809,7257 <sup>a</sup>                  | 7,6685 <sup>a</sup>  | 93,0745 <sup>a</sup>                   | 118,5571 <sup>b</sup>               |
| Desvio padrão         | 0,5029                                 | 0,2792               | 0,1886                                 | 32,7346                             |
| Região “centro-oeste” | Massa específica (kg m <sup>-3</sup> ) | pH                   | Teor alcoólico(g 100 <sup>-1</sup> g)  | Condutividade (µS m <sup>-1</sup> ) |
| Valor mínimo          | 807,7000                               | 6,0000               | 92,5000                                | 60,7000                             |
| Valor máximo          | 811,0000                               | 7,9000               | 93,9000                                | 324,0000                            |
| Valor médio           | 810,0704 <sup>a</sup>                  | 6,6164 <sup>a</sup>  | 92,9085 <sup>a</sup>                   | 136,1897 <sup>a</sup>               |
| Desvio padrão         | 0,6980                                 | 0,439553             | 0,2607                                 | 42,7240                             |

## 5.2 REDES NEURAS DO TIPO PERCEPTRON DE MÚLTIPLAS CAMADAS

O módulo de redes neurais do software Statistica v.9.0, foi configurado para produzir vinte redes e selecionar as cinco com melhor desempenho, a cada inicialização. Foram executadas cinco inicializações independentes, ou seja, sem transmissão de informações e de aprendizado dentre elas. Portanto, no total foram apresentados pelo *software* os dados de 25 redes neurais.

As amostras e parâmetros analisados apresentados às redes foram subdivididos randomicamente em três grupos, um para treinamento abrangendo 70% do total de amostras, outro para teste com 15% das amostras e um para validação com 15% das amostras. As amostras de teste e as de validação não fizeram parte do treinamento e foram utilizadas para a rede aferir sua capacidade de generalização e conferir os resultados obtidos (HAYKIN, 2001; BORSATO et al.,



2009; LINK et al., 2014). A taxa de aprendizagem inicial foi de 0,10, com um número máximo de ciclos de treinamento igual a 200 e de 3 a 10 neurônios selecionados aleatoriamente para a camada oculta.

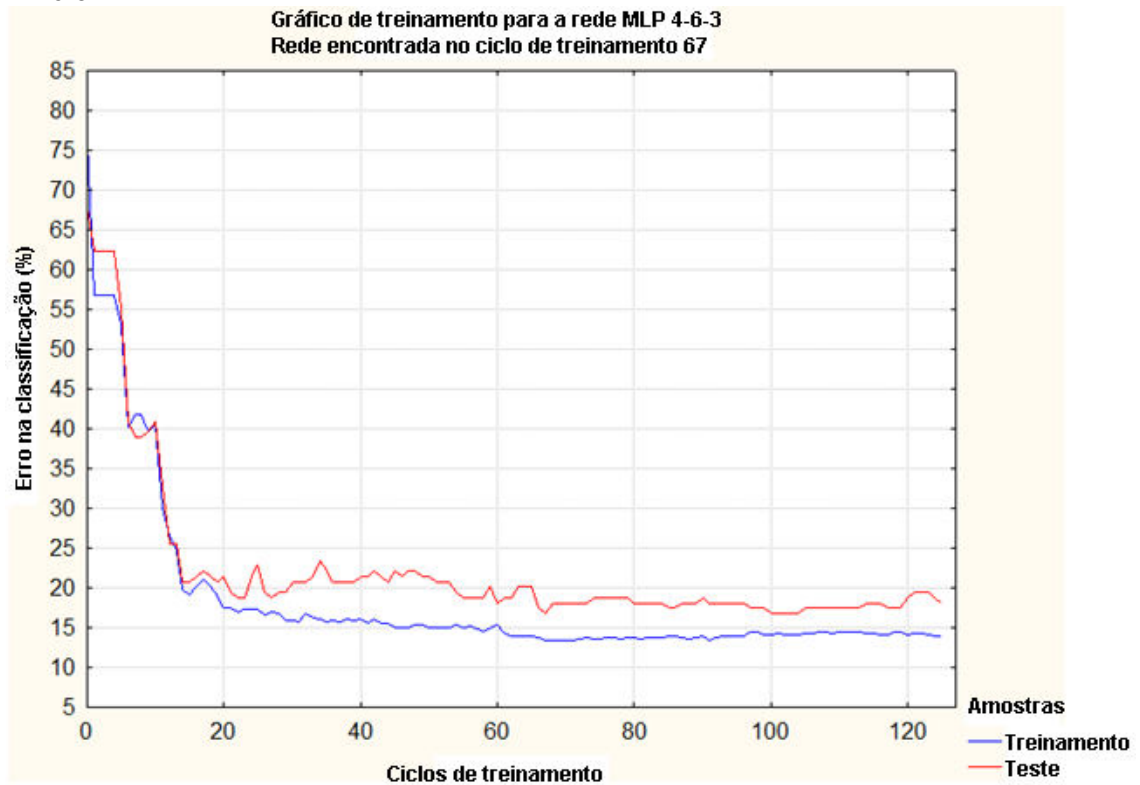
O número de ciclos de treinamento e de neurônios na camada oculta não pode ser muito alto, porque quando uma rede neural aprende muitos exemplos de dados de saída em relação aos de entrada, ela pode acabar memorizando resultados e perder capacidade de generalização. Quanto menor a taxa de aprendizagem, menores serão as variações nos pesos sinápticos de interação e mais suave será a trajetória. Todavia, se a taxa de aprendizagem é muito alta, as modificações irão resultar em maiores pesos sinápticos, que podem tornar a rede instável já que os pesos são ajustados de forma iterativa (HAYKIN, 2001; BISHOP, 2007).

As 25 redes de melhor desempenho, obtiveram acertos com valores aproximados de 75% até 87%, para treinamento, de 78% a 85% para teste e 81% a 87% para validação. O valor médio para treinamento foi de 82,68%, para teste foi de 80,16% e para validação foi de 84,91%. Em estudo para classificação geográfica de amostras de café, a percentagem de acerto da rede no treinamento ficou entre 81 e 86% (ANDERSON & SMITH, 2002).

Como cada inicialização selecionou aleatoriamente o algoritmo de treinamento e as funções de ativação, camada oculta e de erro, ocorrem diferenças de desempenho mesmo com as mesmas quantidades de neurônios na camada oculta.

A cada treinamento o erro foi determinado, e essa informação foi utilizada para ajustar os pesos a fim de fazer o erro decair até sua estabilização (Figura 5). O erro cometido na classificação, em percentagem, decai à medida que o número de ciclos aumenta estabilizando-se a partir do 67º ciclo de treinamento.

Figura 5 – Representação do erro obtido em relação ao número de ciclos de treinamento da rede MLP 4-6-3.



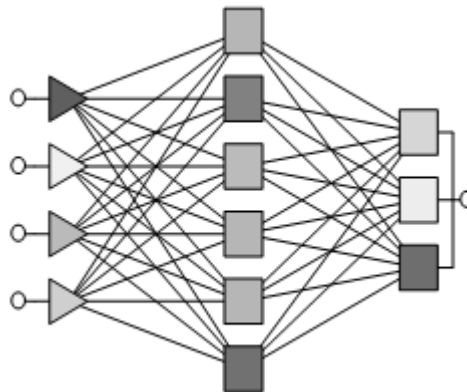
Como cada inicialização foi realizada sem influência da anterior, cada rede é resultado de um aprendizado independente sobre os mesmos dados (Tabela 2). Dentre todas, a rede com melhor desempenho para o treinamento (86,57%) foi estruturada com 6 neurônios na camada oculta, e seus acertos para teste e validação foram respectivamente de 83,22% e 85,23%.

Tabela 2 – Taxa de acerto, e classe de funções utilizadas para as cinco melhores redes neurais obtidas.

| Denominação da rede | Configuração da rede | Treinamento (%) | Teste (%) | Validação (%) | Algoritmo de treinamento | Função de erro | Função da camada oculta | Função de saída |
|---------------------|----------------------|-----------------|-----------|---------------|--------------------------|----------------|-------------------------|-----------------|
| Rede 01             | MLP 4-9-3            | 83,29           | 79,87     | 85,91         | BFGS                     | Entropy        | Exponential             | Softmax         |
| Rede 02             | MLP 4-6-3            | 86,57           | 83,22     | 85,23         | BFGS                     | Entropy        | Logistic                | Softmax         |
| Rede 03             | MLP 4-3-3            | 84,71           | 79,87     | 85,91         | BFGS                     | SOS            | Logistic                | Softmax         |
| Rede 04             | MLP 4-3-3            | 84,71           | 79,19     | 87,25         | BFGS                     | Entropy        | Exponential             | Softmax         |
| Rede 05             | MLP 4-8-3            | 84,29           | 81,21     | 86,58         | BFGS                     | SOS            | Exponential             | Exponential     |

Ao fim dos ciclos de treinamento para aprendizagem de todos os dados, a rede obteve as informações da atividade dos neurônios para cada uma das 998 amostras. Para cada uma a quantidade de neurônios utilizada e a estrutura da rede é a mesma, mas o que as diferencia são as atividades dos neurônios. A Figura 6 é uma representação da Rede 02, com quatro neurônios de entrada representando os parâmetros do etanol, seis na camada oculta e três na camada de saída correspondendo às regiões de comercialização do etanol, sendo que os representados com cor mais escura estão mais ativos. Portanto, o neurônio mais escuro na saída indica a região que a rede relacionou com os parâmetros de entrada daquela amostra de etanol.

Figura 6 – Representação da estrutura rede 02, MLP 4-6-3.



A rede, considerando todas as amostras analisadas, demonstra quais os parâmetros conforme suas relevâncias podem ser utilizadas para a classificação. A Tabela 3 apresenta a análise sensitiva das cinco principais redes. Os parâmetros receberam uma classificação conforme sua ordem de importância, sendo que o número 1 corresponde ao mais importante. A massa específica foi o parâmetro de maior relevância, sendo o mais importante para três redes e o segundo para outras duas. O teor alcoólico é o segundo parâmetro mais relevante, assim considerado por três redes. A Rede 02, que teve o melhor desempenho, corrobora essa classificação colocando como mais importante a massa específica e segundo o teor alcoólico. A condutividade foi o parâmetro menos importante para as cinco redes.

Tabela 3 – Classificação de importância dos parâmetros conforme análise sensitiva.

|         | Massa<br>específica | Teor<br>alcoólico | pH | Condutividade |
|---------|---------------------|-------------------|----|---------------|
| Rede 01 | 1                   | 2                 | 3  | 4             |
| Rede 02 | 1                   | 2                 | 3  | 4             |
| Rede 03 | 2                   | 3                 | 1  | 4             |
| Rede 04 | 1                   | 2                 | 3  | 4             |
| Rede 05 | 2                   | 1                 | 3  | 4             |

A massa específica está correlacionada com o teor alcoólico (Figura 4), o que explica a maior importância de ambas na análise sensitiva.

A rede perceptron com seis neurônios na camada oculta, com o algoritmo de treinamento BFGS (Broyden-Fletcher-Goldfarb-Shanno), função de erro entropia, função de ativação da camada oculta logística e função de saída softmax, foi repetida quarenta vezes para verificação da estabilidade da rede selecionada, sendo que o melhor resultado para treinamento atingiu 86,14%, e o menor 79,29%. Para o teste os resultados variaram de 77,18 a 82,55% e para a validação de 79,19 a 87,25%.

As redes MLP mostraram-se úteis para a classificação de amostras de etanol conforme sua região de comercialização. A análise sensitiva revelou que a massa específica foi a variável mais importante e em seguida o teor alcoólico, sendo que a condutividade foi a menos importante.

### 5.3 REDES NEURAIS DO TIPO MAPAS AUTO-ORGANIZÁVEIS

Os dados de todas as amostras foram utilizados para alimentar o módulo de redes neurais artificiais do tipo mapa auto-organizável (MAO), disponível na *toolbox* de redes neurais do software Matlab R2007, para a segmentação das amostras de álcool etílico combustível de acordo com a região de origem. A rede MAO transformou um padrão de sinal incidente de dimensão arbitrária em um mapa discreto bidimensional, apresentando esta transformação de uma maneira topologicamente ordenada (KOHONEN, 1997).

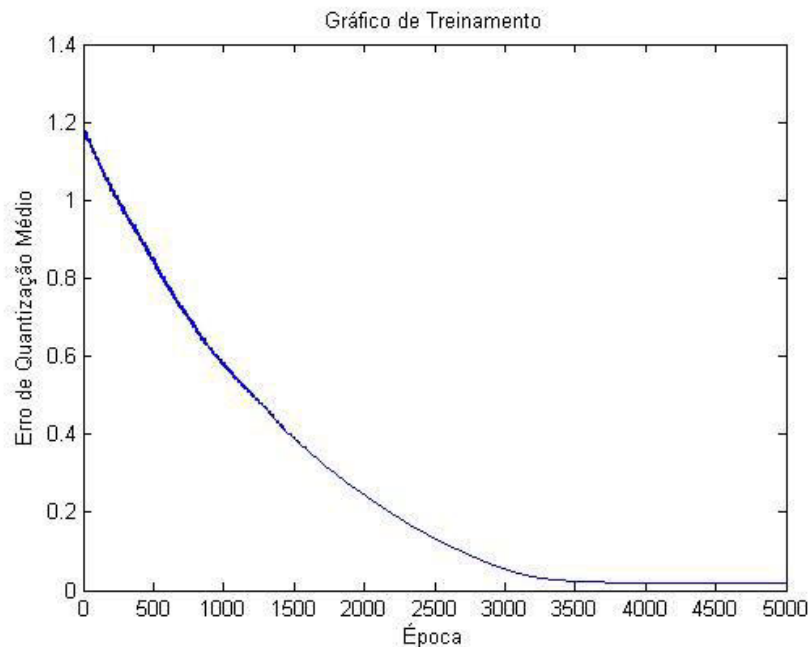
Uma rede foi treinada com topologia hexagonal 35x35, 7000 épocas de treinamento, com taxa de aprendizagem inicial de 0,1 e relação de vizinhança inicial de 17. A rede estabilizou o erro de quantização médio em  $5 \times 10^{-2}$  com 5000 épocas

de treinamento. A taxa de aprendizagem decaiu até  $9,1 \times 10^{-5}$  e o valor final da relação de vizinhança foi de 0,003. Nesta topologia ocorreram 50 casos em que mais de uma amostra de regiões de origem diferentes ocupam o mesmo neurônio, por sobreposição, o que significa que a malha de neurônios não foi suficiente para possibilitar a separação de grupos.

De acordo com Boisherebert et al. (2006), a topologia deve ser escolhida de modo que não possua poucos neurônios em relação ao número das amostras, para que estas não se sobreponham. Entretanto, não pode ter um número de neurônios muito grande para que as amostras não se dispersem exageradamente, pois assim podem ser formados mais grupos do que desejado. Além disso, uma topologia maior provocaria desperdício de tempo de processamento computacional.

Portanto, uma rede de topologia hexagonal  $45 \times 45$  foi treinada, com 5000 épocas de treinamento, com taxa de aprendizagem inicial de 0,1 e relação de vizinhança inicial de 22. A rede estabilizou o erro em 4000 épocas (Figura 7). O valor final do erro foi de 0,02, a taxa de aprendizagem decaiu até  $6.7 \times 10^{-4}$  e o valor final da relação de vizinhança foi de 0,03.

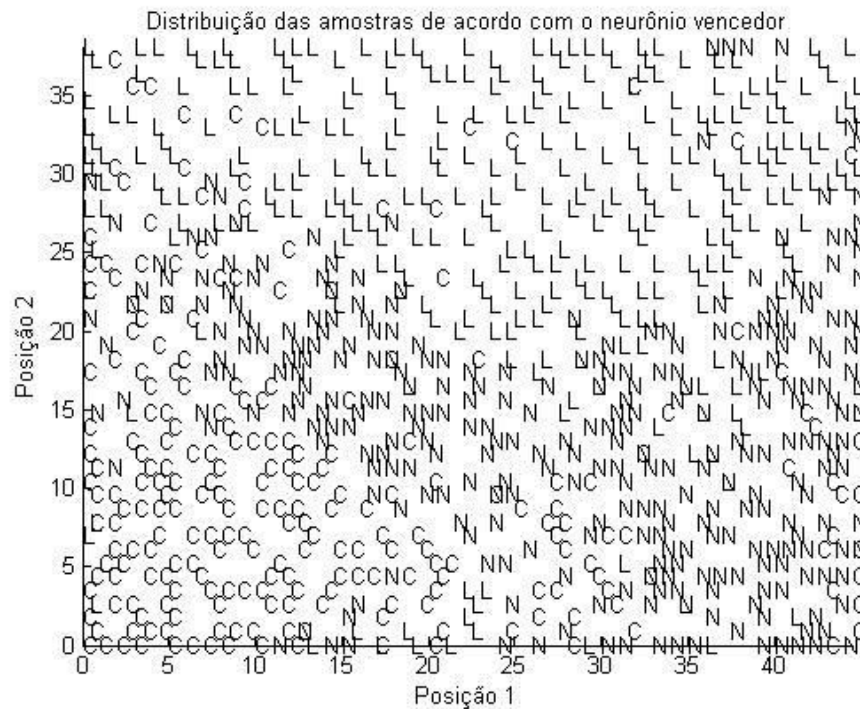
Figura 7 – Erro de quantização da rede MAO em relação ao número de épocas.



Conforme distribuição na topologia  $45 \times 45$ , pode-se observar que as amostras da região “leste” (L) localizam-se na parte superior, as do “norte” (N) na

parte intermediária e as do “centro-oeste” (C) na inferior (Figura 8). As amostras foram identificadas de acordo com o neurônio vencedor para cada uma, buscando-se separá-las o máximo possível em diferentes áreas do mapa, a partir de suas características inerentes. Portanto, três grupos podem ser visualmente identificados, tendo ocorrido apenas 26 casos de sobreposição de amostras de regiões diferentes no mesmo neurônio indicando que a rede utilizada apresentou um menor erro de classificação em relação à rede anterior treinada de topologia 35x35.

Figura 8 – Gráfico de treinamento para a rede MAO de configuração 45x45, contendo a distribuição das amostras das regiões “norte” (N), “leste” (L) e “centro-oeste” (C).

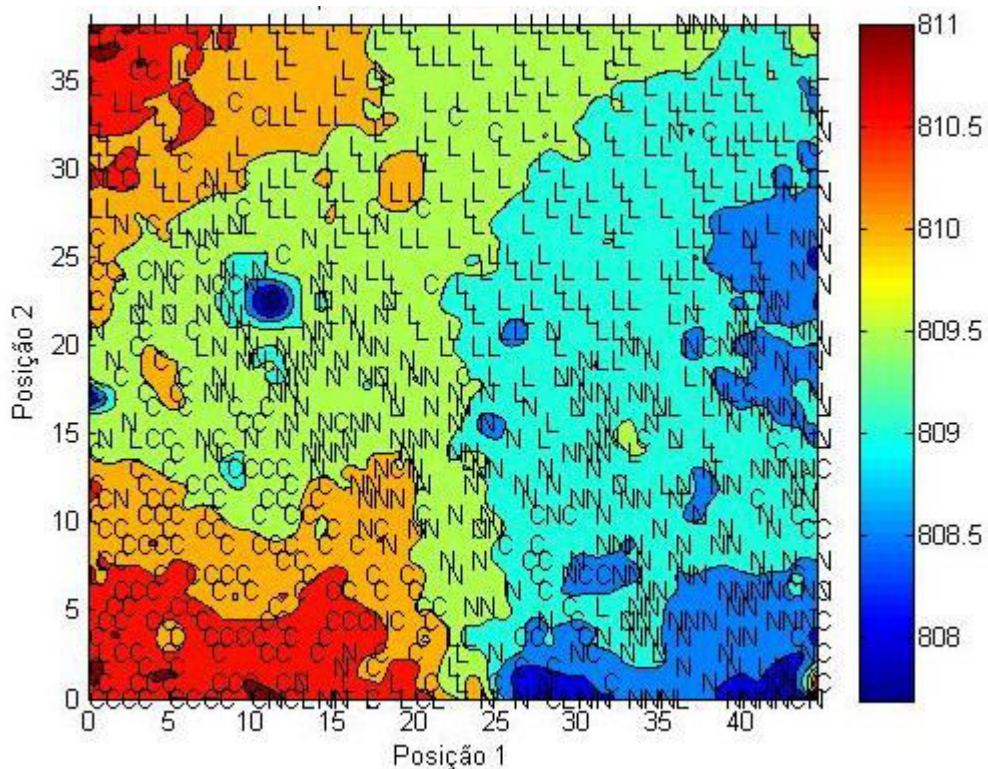


A configuração do mapa topológico (Figura 8) formou-se considerando quatro variáveis de entrada, equivalentes aos parâmetros analisados nas amostras de etanol. Sobre esse mapa podem ser sobrepostas faixas de pesos referentes a cada variável, gerando-se um mapa de pesos para cada parâmetro. As amostras podem ser identificadas dentro de faixas de pesos, designadas por tonalidades diferentes, sendo que amostras pertencentes à mesma faixa possuem pesos semelhantes para aquela variável. A presença de amostras de etanol, de mesma região de origem de comercialização, predominando em faixas de pesos iguais ou próximas significa que aquele parâmetro foi significativo para segmentar aquelas

amostras das demais. Apesar de cada mapa apresentar os resultados de somente uma variável, as influências de todas as demais foram consideradas no seu cálculo.

O mapa de pesos para a massa específica (Figura 9) indica que as amostras da região “centro-oeste” predominam nas cores vermelha e alaranjada, com pesos maiores, diferenciando-a das amostras das regiões “norte” e “leste”. Porém, não foi possível fazer distinção entre as amostras das regiões “leste” e “norte”, pois apresentam resultados em faixas de pesos semelhantes, dispostas das regiões azuis do mapa. Portanto, a massa específica foi um parâmetro importante somente para diferenciar as amostras do “centro-oeste” das demais.

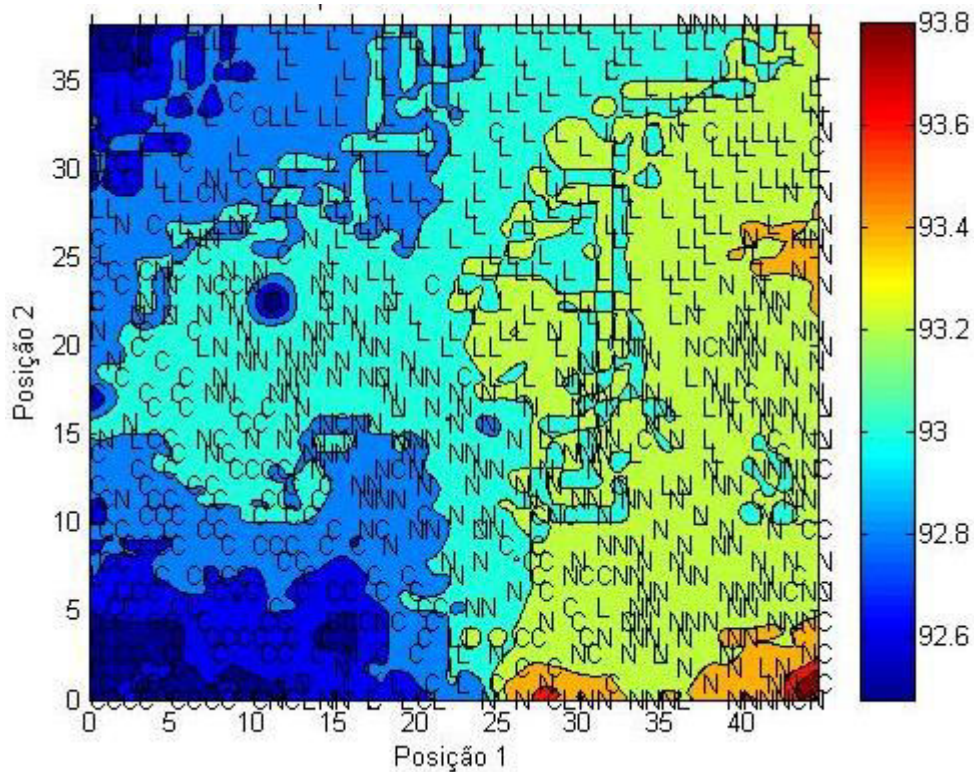
Figura 9 – Mapa de pesos para a massa específica



A variável teor alcoólico, também foi importante para a separação das amostras do “centro-oeste”, pois apesar de existirem amostras do grupo “leste” situadas em pesos menores, a maioria do grupo “centro-oeste” está localizado nessas regiões, ou seja, nas áreas de cor azul (Figura 10). Já as amostras da região “norte” e “leste” encontram-se distribuídas de forma heterogênea no mapa, em todas as faixas de pesos.

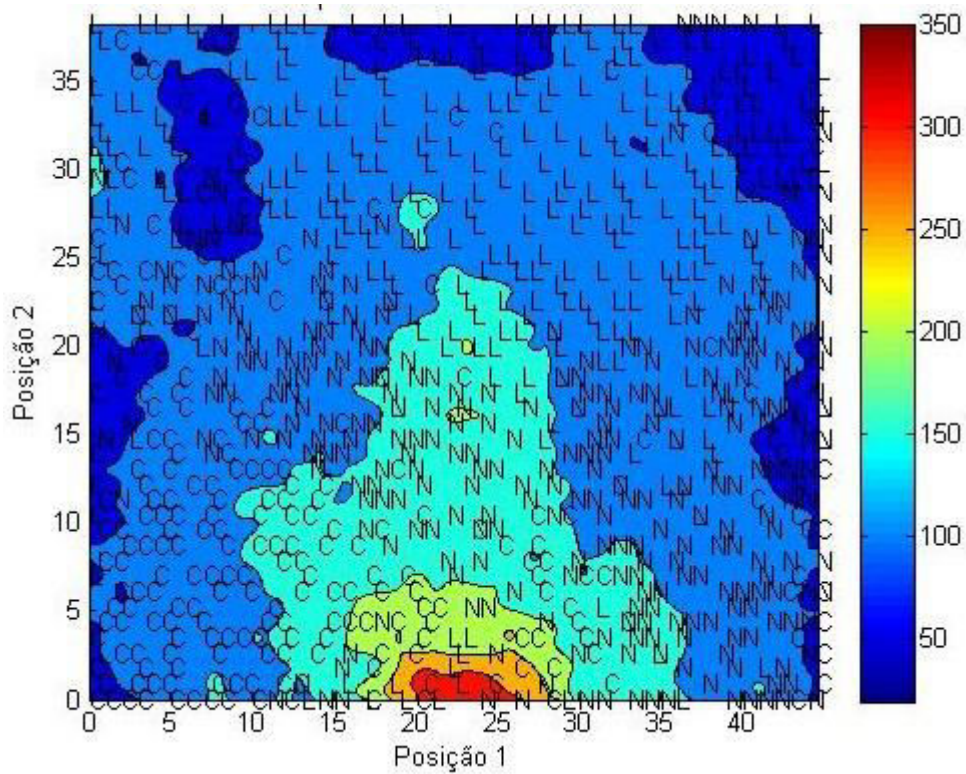


Figura 10 – Mapa de pesos para o teor alcoólico



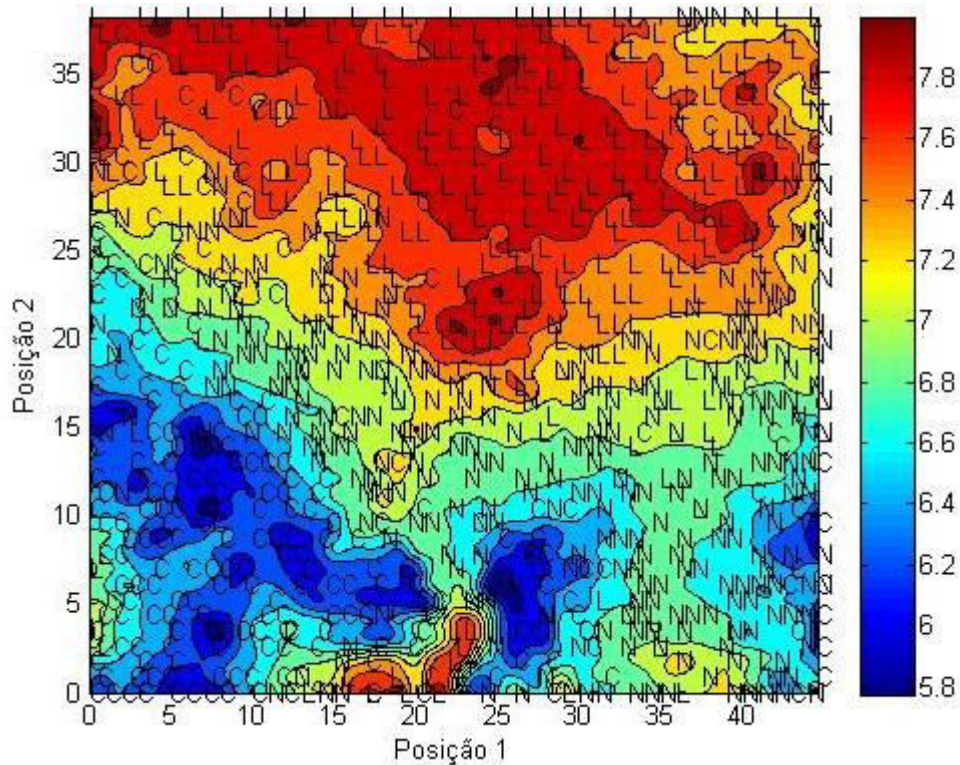
A condutividade elétrica apresenta amostras de todas as regiões em todas as faixas de pesos (Figura 11). Apenas há tendência das amostras do “centro-oeste” e do “norte” a estarem nas regiões vermelha e amarela, de pesos maiores. Nas áreas do mapa de cor azul, correspondentes a pesos menores, existem regiões que tendem a concentrar amostras do “leste”.

Figura 11 – Mapa de pesos para a condutividade elétrica.



Para variável pH, o grupo “leste” está situado predominantemente nas regiões de cor vermelha e alaranjada do mapa, que corresponde a pesos maiores (Figura 12). O grupo “centro-oeste” está situado nas regiões azuis do mapa, ou seja, de pesos menores. O grupo “norte” está localizado em uma região de pesos intermediários. Além de indicar a formação dos três grupos, esse parâmetro também foi importante para justificar o distanciamento dos grupos “leste” e “centro oeste”, por estes se encontrarem em regiões de pesos muito diferentes.

Figura 12 – Mapa de pesos para o pH



Os resultados, portanto, mostram que foi possível distinguir amostras de álcool etílico hidratado combustível, comercializadas nas regiões “norte”, “centro-oeste” e “leste” do Paraná, através da utilização de redes neurais do tipo mapas auto-organizáveis. A utilização da topologia 45x45 foi a que apresentou melhor separação das amostras, sendo suficientes 4000 épocas de treinamento da rede. Os parâmetros mais influentes na diferenciação foram a massa específica e o pH e, em seguida, o teor alcoólico. A condutividade apresentou amostras de todas as regiões em todas as faixas do seu mapa de pesos, apenas indicando tendências de agrupamentos, sendo menos relevante que os demais parâmetros.

## 6 CONCLUSÃO

As redes neurais artificiais do tipo perceptron de múltiplas camadas (MLP) e mapas auto-organizáveis (MAO), aplicadas para analisarem dados físico-químicos de 998 amostras de etanol hidratado combustível, mostraram-se como ferramentas úteis para o reconhecimento de padrões significativos que possam diferenciar as amostras. As redes neurais puderam reconhecer que amostras de etanol possuem diferenças de acordo com a sua região de comercialização.

No que diz respeito às vantagens e desvantagens de cada tipo de rede neural, as redes MLP possuíram tempo de operação computacional de alguns minutos, enquanto que as redes MAO necessitaram de dezenas de horas de processamento. Contudo, as redes MAO apresentaram resultados que podem ser verificados de forma visual. Destaca-se que cada tipo de rede mostrou estabilidade quanto a seus resultados quando da repetição dos experimentos.

Ambos os tipos de redes neurais foram capazes de executar análise sensível apontando quais parâmetros físico-químicos exercem maior influência numa classificação de amostras. Além disso, apresentaram resultados concordantes quanto à relevância de cada parâmetro de conformidade na segmentação das amostras, sendo a condutividade elétrica o menos importante e a massa específica a 20 °C o mais importante.

## **7 SUGESTÕES PARA TRABALHOS FUTUROS**

- Aplicação de redes neurais para classificação de amostras de etanol conforme o tipo de matéria-prima adotada para sua produção.
- Aplicação de redes neurais para segmentar amostras de etanol de acordo com seus diferentes processos produtivos.
- Análise sensitiva de variáveis, através do uso de redes neurais, para estudar a importância de parâmetros de análises físico-químicas (variáveis de entrada) em amostras de etanol de categorias diversas.

## REFERÊNCIAS

- AGARWAL, A. K. Biofuels (alcohols and biodiesel) applications as fuels for internal combustion engines. **Progress in Energy and Combustion Science**, v. 33, pp. 233–271, 2007.
- ANDERSON, K. A., & SMITH, B. W. Chemical Profiling To Differentiate Geographic Growing Origins of Coffee. **Journal of Agricultural and Food Chemistry**, v. 50, n. 7, pp. 2068-2075, 2002.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 5992**: Álcool etílico e suas misturas com água - Determinação da massa específica e do teor alcoólico - Método do densímetro de vidro. Rio de Janeiro, 2009, 5 p.
- \_\_\_\_\_. **NBR 10891**: Etanol hidratado combustível - Determinação do pH - Método potenciométrico. Rio de Janeiro, 2013, 3 p.
- \_\_\_\_\_. **NBR 10547**: Etanol combustível - Determinação da condutividade elétrica. Rio de Janeiro, 2013, 3 p.
- BALLERINI, D.; ALAZARD-TOUX, N. **Les biocarburants: état des lieux, perspectives et enjeux du développement**. Paris: Éditions Technip, 2006. 348p. ISBN 978-2-7108-0869-5.
- BASTOS, V. D. Etanol, Alcoolquímica e Biorrefinarias, **BNDES Setorial**, Rio de Janeiro, n. 25, p. 5-38, mar. 2007.
- BISHOP, C. M. **Neural Networks for Pattern Recognition**. New York: Oxford University Press, 2007.
- BOISHEBERT, V.; GIRAUDEL, J. L.; MONTURY, M. Characterization of strawberry varieties by SPME-CG-MS and Kohonen self-organizing map, **Chemometrics and Intelligent Laboratory Systems**, v. 80, pp.13-23, 2006.
- BONACHERA, F.; MARCOU, G.; KIREEVA, N.; VARNEK, A.; HORVATH, D. Using self-organizing maps to accelerate similarity search, **Bioorganic & Medicinal Chemistry**, v. 20, pp. 5396–5409, 2012.
- BORSATO, D., MOREIRA, I., NOBREGA, M. M., MOREIRA, M. B., DIAS, G. H., SILVA, R. S. S. F., & BONA, E. Aplicação de Redes Neurais Artificiais na Identificação de Gasolinas Adulteradas Comercializadas na Região de Londrina – Paraná, **Química Nova (Impresso)**, v. 32, n. 9, pp. 2328-2332, 2009.
- BORSATO, D., PINA, M. V. R., SPACINO, K. R., SCHOLZ, M. B. S., & ANDROCIOLI, A. Application of Artificial Neural Networks in the Geographical Identification of Coffee Samples, **European Food Research & Technology**, v. 233, n. 3, pp. 533-543, 2011.

BRASIL. Agência Nacional de Petróleo, Gás Natural. Resolução n° 7, de 9 de fevereiro de 2011. Estabelece a especificação para a comercialização do álcool etílico hidratado combustível. **Diário Oficial da União**, Brasília, DF, 14 abr 2011.

\_\_\_\_\_. Agência Nacional de Petróleo. Resolução n° 7, de 21 de fevereiro de 2013. Estabelece a especificação para a comercialização do álcool etílico hidratado combustível. **Diário Oficial da União**, Brasília, DF, 15 abr 2013.

\_\_\_\_\_. Agência Nacional de Petróleo, Gás Natural e Biocombustíveis. Resolução n° 19, de 15 de abril de 2015. Estabelece a especificação para a comercialização do álcool etílico hidratado combustível. **Diário Oficial da União**, Brasília, DF, 20 abr 2015.

\_\_\_\_\_. Agência Nacional de Petróleo, Gás Natural e Biocombustíveis. Monitoramento da Qualidade de Combustíveis e Lubrificantes. 2016a Disponível em: <<http://www.anp.gov.br/?pg=46607>>. Acesso em: 21 mar 2016.

\_\_\_\_\_. Agência Nacional de Petróleo, Gás Natural e Biocombustíveis. PQMC - Boletim de monitoramento da qualidade dos combustíveis. 2016b. Disponível em: <<http://www.anp.gov.br/?pg=79876&m=pmqc&t1=&t2=pmqc&t3=&t4=&ar=0&ps=1&1458570913076>>. Acesso em: 21 mar 2016.

BURTON, E. J.; SCHOLEY, D. V.; WILLIAMS, P. E. V. Use of cereal crops for food and fuel – characterization of a novel bioethanol coproduct for use in meat poultry diets, **Food and Energy Security**, v. 2, n. 3, pp. 197–206, 2013.

CARNEIRO, H. S. P.; MEDEIROS, A. R. B.; OLIVEIRA, F. C. C.; AGUIAR, G. H. M.; RUBIM, J. C.; SUAREZ, P. A. Z. Determination of Ethanol Fuel Adulteration by Methanol Using Partial Least-Squares Models Based on Fourier Transform Techniques, **Energy & Fuels**, 2008, v. 22, pp. 2767–2770, 2008.

CAETANO, L. G.; TAKEUCHI, R. M.; SANTOS, A. L.; OLIVEIRA, M. F. O.; STRADIOTTO, N. R. Voltammetric determination of ethyl acetate in ethanol fuel using a Fe<sup>3+</sup>/Nafion®-coated glassy carbon electrode, **Fuel**, v. 106, pp. 837–842, 2013.

CREMASCO, H., BORSATO, D., ANGILELLI, K. G., GALÃO, O. F., BONA, E., VALLE, M. E. Application of self-organising maps towards segmentation of soybean samples by determination of inorganic compounds content, **J Sci Food Agric**, v. 96, n.1, pp. 306-310, 2016.

DALMOLIN, I.; SKOVROINSKI, E.; BIASI, A.; CORAZZA, M. L.; DARIVA, C.; OLIVEIRA, J. V. Solubility of carbon dioxide in binary and ternary mixtures with ethanol and water. **Fluid Phase Equilibria**, v. 245, pp. 193-198, 2006.

DELGADO, R.; ARAUJO, A. S.; FERNANDES Jr, V. J. Properties of Brazilian gasoline mixed with hydrated ethanol for flex-fuel technology, **Fuel Processing Technology**, v. 88, pp. 365–368, 2007.

DEISINGH, A. K.; STONE, D. C.; THOMPSON, M. Applications of electronic noses and tongues in food analysis, **International Journal of Food Science and Technology**, v. 39, n. 6, p. 587-604, 2004.

EBADZADEH, M.; DARLOT, C. Cerebellar Learning of Bio-Mechanical Functions of Extra-Ocular Muscles: Modeling by Artificial Neural Networks, **Neuroscience**, v. 122, pp. 941-966, 2003.

FELIPE, M. G. A. Bioetanol de Cana-de-Açúcar: P&D para Produtividade Sustentabilidade, Cortez, L. A. B., ed.; Edgard Blücher Ltda.: São Paulo, cap. 3, parte 4, 2010.

FIORIN, D. V.; MARTINS, F. R.; SCHUCH, N. J.; PEREIRA, E. B. Aplicações de redes neurais e previsões de disponibilidade de recursos energéticos solares, **Revista Brasileira de Ensino de Física**, v. 33, n. 1, 1309, 2011.

GEACH, J. E. Unsupervised self-organized mapping: a versatile empirical tool for object selection, classification and redshift estimation in large surveys, **Mon. Not. R. Astron. Soc.**, v. 419, pp. 2633–2645, 2012.

GONÇALVES, M. A.; GONZAGA, F. B.; FRAGA, I. C. S.; RIBEIRO, C. M.; SOBRAL, S. P.; REGO, E. C. P.; SANTANA, E. B.; OLIVEIRA, L. M.; SILVA, V. F.; LEAL, R. V. P.; SANTOS FILHO, D. M. E.; SIQUEIRA, J. R. R.; BARBOSA, T. P.; RODRIGUES, J. M.; CUNHA, V. S.; SKROBOT, V. L.; COSTA, C. B.; PESSOA Jr., A.; CARNEIRO, H. S. P.; COLARES, H.; CALIMAN, E.; ALVES, M. V. S. Avaliação de laboratórios brasileiros na determinação de alguns parâmetros de qualidade de biocombustíveis. **Química Nova**, v. 36, n.º. 3, pp. 393-399, 2013.

GONÇALVES, M. A.; GONZAGA, F. B.; FRAGA, I. C. S.; RIBEIRO, C. M.; SOBRAL, S. P.; BORGES, P. P.; ROCHA, W. F. C. Evaluation study of different glass electrodes by an interlaboratory comparison for determining the pH of fuel ethanol, **Sensors and Actuators B: Chemical**, v. 158, pp. 327– 332, 2011.

GONZAGA, F. B.; GONÇALVES, M. A.; SOBRAL, S. P.; RIBEIRO, C. M. A new method for determining the acid number of ethanol fuel using coulometric titration, **Fuel**, v. 94, pp. 70–74, 2012.

HAIR Jr., J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. Análise Fatorial. In **Análise Multivariada de Dados**. 6ª. ed., Bookman: Porto Alegre, cap.3, pp. 89-127, 2009.

HAYKIN, S. **Redes Neurais: Princípio e Práticas**, 2ª. ed., Bookman: Porto Alegre, 2001.

HAXBY, J. V.; CONNOLLY, A. C.; GUNTUPALLI, J. S. Decoding Neural Representational Spaces Using Multivariate Pattern Analysis. **Annual Review of Neuroscience**, v. 37, pp. 435-456, 2014.



IGLESIAS, C. & SESMERO, J. P. Economic Analysis of Supplementing Sugarcane with Corn for Ethanol Production in Brazil: A Case Study in Uberaba. **Bioenerg. Res.**, v. 8, pp. 627–643, 2015.

JOANNIS-CASSAN, C.; RIESS, J.; JOLIBERT, F.; TAILLANDIER, P. Optimization of very high gravity fermentation process for ethanol production from industrial sugar beet syrup. **Biomass and Bioenergy**, v. 70, pp.165-173, 2014.

KAMEI, N.; KIKUCHI, S.; TAKEDA-MORISHITA, M.; TERASAWA, Y.; YASUDA, A.; YAMAMOTO, S.; IDA, N.; NISHIO, R.; TAKAYAMA, K. Determination of the Optimal Cell-Penetrating Peptide Sequence for Intestinal Insulin Delivery Based on Molecular Orbital Analysis with Self-Organizing Maps. **Journal of Pharmaceutical Sciences**, v. 102, n. 2, pp. 469-479, 2013.

KOHONEN, T. **Self Organizing Maps**. Series in Information Sciences. v. 30, 2<sup>a</sup>. ed, Editions Springer – Verlag, Heidelberg, 1997.

LA HARPE, E. R. Ignition-improved ethanol as a diesel tractor fuel. Dissertation (M. Sc. Eng.) - Department of Agricultural Engineering. University of Natal Pietermaritzburg, Durban-África do Sul, 1988.

LIAO, S. Expert system methodologies and applications – a decade review from 1995 to 2004, **Expert Systems with Applications**, v.28, pp.93-109, 2005.

LINK, J. V.; LEMES, A. L. G.; MARQUETTI, I.; SCHOLZ, M. B. S.; BONA, E. Geographical and genotypic segmentation of arabica coffee using self-organizing maps. **Food Research International**, v.59, pp. 1–7, 2014.

LOUIS, P.; SERET, A.; BAESENS, B. Financial Efficiency and Social Impact of Microfinance Institutions Using Self-Organizing Maps. **World Development**, v. 46, pp. 197–210, 2013.

LUCCHINI, M.; ASSI, J. Mapping Patterns of Multiple Deprivation and Well-Being using Self-Organizing Maps: An Application to Swiss Household Panel Data, **Social Indicators Research**, v. 112, pp. 129–149, 2013.

LYRA, W. S.; SILVA, E. C. DA; ARAÚJO, M. C. U.; FRAGOSO W. D. Classificação periódica: um exemplo didático para ensinar análise de componentes principais. **Química Nova**, v. 33, n. 3, pp. 1594-1597, 2010.

OZAKI, S. T. R. Detecção de adulteração de combustíveis com sensores poliméricos eletrodepositados e redes neurais artificiais. 2010. Dissertação (Mestrado em Engenharia Elétrica) - Escola Politécnica da Universidade de São Paulo, São Paulo, 2010.

RIBEIRO, F. A. L.; MANTOVANI, G. A.; POPPI, R. J.; ROSÁRIO, F. F.; BEZERRA, M. C. M.; BASTOS, A.L.M.; MELO, V. L. A. PCA: uma ferramenta para identificação de traçadores químicos para água de formação e água de injeção associadas à produção de petróleo, **Química Nova**, v. 36, n. 9, pp. 1281-1287, 2013.

MATLAB. Neural Network Toolbox. The MathWorks Inc. Versão R2007.

MELLER, C.; GENTER, A.; KOHL, T. The application of a neural network to map clay zones in crystalline rock, **Geophysical Journal International**, v. 196, n. 2, pp. 837-849, 2014.

OLIVEIRA, M. F.; SACK, A. A.; OKUMURA L. L.; STRADIOTTO, N. R. Analytical Methods Employed at Quality Control of Fuel Ethanol, **Energy Fuels**, v. 23, pp. 4852–4859, 2009.

OWEN, K.; COLEY, T. **Automotive Fuels Reference Book**. 2<sup>a</sup>. ed., Society of Automotive Engineers: Warrendale, 1995.

OZAKI, S. T. R. **Detecção de adulteração de combustíveis com sensores poliméricos eletrodepositados e redes neurais artificiais**. 2010, USP.

RENDE, S; DONDURAN, M. Neighborhoods in Development: Human Development Index and Self-organizing Maps, **Social Indicators Research**, v. 110, pp. 721–734, 2013.

RITTER, H. Self-organizing feature maps: Kohonen maps. In ARBIB, M. A. **The Handbook of Brain Theory and Neural Network**. Cambridge: MIT Press, pp. 846-851, 1995.

ROMAGNOLI, E.; SILVA, L. R. C.; ANGILELLI, K. G.; FERREIRA, B. A. D.; WALKOFF, A. R.; BORSATO, D. Using Neural Network Multilayer Perceptron type for segmentation samples of hydrated ethanol by commercialisation region, **Acta Scientiarum**, 2016 (in press).

SIERRA, F.; GUILLOU, C. G.; RENIERO, F.; BALLARIN, L.; CANTAGALLO M. I.; WIESER, M.; IYER, S. S.; HÉBEGER, K.; VANHAECKE, F.; Determination of the geographical origin of green coffee by principal component analysis of carbon, nitrogen and boron stable isotope ratios, **Rapid Communications in Mass Spectrometry**, v. 19, p. 2111-2115, 2005.

SILVA, J. A.; SILVA, F. L. H.; ALVES, R. R. N.; SANTANA, D. P. Influência das variáveis nitrogênio, fósforo e o °Brix na produção dos metabólitos secundários contaminantes totais da fermentação alcoólica, **Química Nova**, v. 29, n.4, pp. 695-698, 2006.

SILVA, A. C.; PAZ, J. E. M.; PONTES, L. F. B. L.; LEMOS, S. G.; PONTES, M. J. C. An electroanalytical method to detect adulteration of ethanol fuel by using multivariate analysis. **Electrochimica Acta**, v.111, pp. 160– 164, 2013.

SILVA, J. A.; DAMASCENO, B. P. G. L.; SILVA, F. L. H.; MADRUGA, M. S.; SANTANA, D. P. Aplicação da metodologia de planejamento fatorial e análise de superfícies de resposta para otimização da fermentação alcoólica, **Química Nova**, v. 31, n. 5, pp. 1073-1077, 2008.

SILVA, L. R. C., ANGILELLI, K. G., CREMASCO, H., ROMAGNOLI, E. S., WALKOFF, A. R., BORSATO, D., Application Self-organizing Map Type in a Study of the Profile of Gasoline C Commercialized in the Eastern and Northern Parana Regions, **Orbital: The Electronic Journal of Chemistry**, v. 7, n. 2, pp. 185-190, 2015.

SMITH Jr., W. O.; AINLEY, D. G.; ARRIGO, K. R.; DINNIMAN, M. S. et al.. The Oceanography and Ecology of the Ross Sea., **Annual Review of Marine Science**, v. 6, pp. 469-487, 2013.

SPACINO, K.; SILVA, H.; ANGILELLI, K.; SILVA, E.; MOREIRA, I.; MESQUITA, M.; BORSATO, D. Using Self-organizing Maps as a Chemometric Tool for Alcohol Classification by Distillery, **International Journal of Environment and Bioenergy**, v. 8, n. 1, pp. 1-11, 2013.

STATISTICA.Graphics software. Tulsa: Statistica for Windows, versão 9, 2009.

TANAKA, G. T.; FERREIRA, F. O.; FERREIRA DA SILVA, C. E.; FLUMIGNAN, D. L.; OLIVEIRA, J. E. Chemometrics in fuel science: demonstration of the feasibility of chemometrics analyses applied to physicochemical parameters to screen solvent tracers in Brazilian commercial gasoline, **Journal of Chemometrics**, v. 25, n. 9, pp. 487-495, 2011.

VANDERHAEGEN, F.; ZIEBA, S. Reinforced learning systems based on merged and cumulative knowledge to predict human actions, **Information Sciences**, v. 276, pp.146-159, 2014.

WANG, F.; JIANG, Y.; GUO, W.; NIU, K.; ZHANG, R.; HOU, S.; WANG, M.; Yi, Y; ZHU, C.; JIA, C.; FANG, X. An environmentally friendly and productive process for bioethanol production from potato waste. **Biotechnology for Biofuels**, v. 9, iss 50, 2016.