

**UNIVERSIDADE ESTADUAL DE CAMPINAS**  
**FACULDADE DE ENGENHARIA QUÍMICA**

**ÁREA DE CONCENTRAÇÃO**  
**SISTEMAS DE PROCESSOS QUÍMICOS E INFORMÁTICA**

**“PREDIÇÃO DO EQUILÍBRIO LÍQUIDO-VAPOR ATRAVÉS  
DE REDES NEURAS ARTIFICIAIS”**

**Autora: Sílvia Dal’Boit**

**Orientador: Prof. Dr. Roger Josef Zemp**

**Dissertação de Mestrado apresentada à Faculdade de Engenharia Química  
como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia  
Química.**

**Junho – 2008**

**Campinas – São Paulo – Brasil**

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE -  
UNICAMP

D15p Dal'Boit, Sílvia  
Predição do equilíbrio líquido-vapor através de redes neurais artificiais / Sílvia Dal'Boit.--Campinas, SP: [s.n.], 2008.

Orientador: Roger Josef Zemp.  
Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Química.

1. Redes neurais (Computação). 2. Equilíbrio líquido-vapor. 3. Termodinâmica. I. Zemp, Roger Josef. II. Universidade Estadual de Campinas. Faculdade de Engenharia Química. III. Título.

Título em Inglês: Prediction of vapor-liquid equilibrium using artificial neural networks.

Palavras-chave em Inglês: Neural networks ( Computer), Vapor – liquid equilibrium, Thermodynamic.

Área de concentração: Sistemas de Processos Químicos e Informática.

Titulação: Mestre em Engenharia Química

Banca examinadora: Maria Alvina Krähenbühl e Jones Erni Schmitz.

Data da defesa: 23/6/2008.

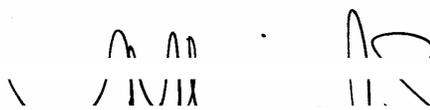
Programa de Pós-Graduação: Engenharia Química.

Dissertação de Mestrado defendida por Sílvia Dal'Boit e aprovada em 23 de junho de 2008 pela banca examinadora constituída pelos doutores:



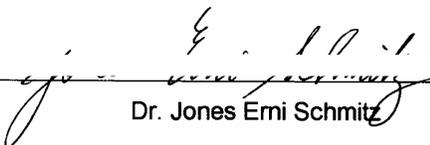
---

Prof. Dr. Roger Josef Zemp - Orientador



---

Prof. Dr. Maria Alvina Krähenbühl



---

Dr. Jones Erni Schmitz

Este exemplar corresponde à versão final da Dissertação de Mestrado em Engenharia Química, defendido por Silvia Dal'Boit em 23 de junho de 2008.



---

Prof. Dr. Roger Josef Zemp

Orientador

## **Agradecimentos**

Agradeço a Deus pelas alegrias, pelas tristezas e por todo o aprendizado ocorrido nestes dois anos de mestrado.

Agradeço a minha família e especialmente a minha mãe, pelo apoio e amor incondicional.

Agradeço ao meu orientador Prof. Dr. Roger Josef Zemp, pela paciência e pela orientação.

Agradeço aos funcionários da FEQ e em especial ao Juliano pela ajuda com os computadores e pelo apoio.

Agradeço aos meus amigos de Campinas, em especial à Francine e à Patrícia pelo apoio, pelas baladas, pelas consultas sentimentais, por terem cuidado de mim sempre que eu precisei e por serem mais do que amigas.

Agradeço à UNICAMP pelas oportunidades, pelas aulas de extensão (dança de salão, tecido circense), pelas festas, pelas pessoas diferentes de cada canto do Brasil e do mundo encontradas aqui, que me proporcionaram esses anos mais que especiais na minha vida.

Agradeço à Capes pelo apoio financeiro.

## Resumo

Esse trabalho tem como objetivo o estudo da viabilidade de se utilizar se redes neurais artificiais (RNAs) para a predição de equilíbrio líquido-vapor (ELV), utilizando a abordagem do método de contribuição de grupos, método UNIFAC. Foram feitas duas propostas de trabalho, uma substituindo o método UNIFAC para o cálculo do coeficiente de atividade (?) pelas RNAs e outra calculando diretamente a composição da fase vapor (y) através das RNAs. A principal vantagem da utilização de uma rede neural reside na sua capacidade de modelar o sistema em equilíbrio sem a necessidade da definição prévia de um modelo de equilíbrio específico, como ocorre, por exemplo, no método UNIFAC.

Foram estudados dois sistemas, um sistema constituído de misturas binárias álcool-água com cinco grupos funcionais UNIFAC e outro com misturas binárias de: álcoois, água, aldeídos e cetonas, com oito grupos funcionais UNIFAC. Foram usadas redes neurais do tipo *feedforward* com uma camada escondida. Utilizou-se o *software* Matlab 7.0 como ferramenta para esse trabalho.

Nesse trabalho o número de neurônios na camada escondida foi estimado por três métodos diferentes. No entanto, os melhores resultados não foram obtidos por nenhum destes e sim por redes com números de neurônios bem menores. Esse resultado demonstrou a independência de teorias e postulados na área de redes neurais artificiais, já que não existe como prever o seu desempenho e nem como escolher através de uma forma ótima e única os seus algoritmos e funções.

Os resultados obtidos mostraram uma boa capacidade de predição da rede, com desvios na mesma ordem de grandeza que os métodos baseados no uso de um modelo específico para a descrição do comportamento de equilíbrio líquido-vapor, mas sem a necessidade de se definir um modelo ou de serem calculados parâmetros de interação ou constantes ajustáveis. Esse resultado é relevante, uma vez que os sistemas estudados apresentam não idealidades acentuadas.

## Abstract

In this work the suitability of artificial neural networks (ANN) for the prediction of vapour-liquid equilibrium (VLE) is analyzed, using molecular groups that describe the components as input. Two different approaches were studied. In the first proposal the ANN is used to predict activity coefficients, in a similar way as done by the UNIFAC method. In the second proposal, vapour composition and equilibrium temperature are computed by the neural network without intermediate activity coefficient calculation. The main gain of the neural network is its capability to model the VLE without a previous definition of a specific equilibrium model, like the UNIFAC model, for instance.

Two different systems were studied, one of these is consisting of only binary mixtures for alcohols-water components with five UNIFAC functional groups. The other one is composed of binary mixtures of alcohols, water, aldehydes and ketones components with eight UNIFAC functional groups. The ANN used in this work is a feedforward network with one hidden layer, the networks were implemented in Matlab 7.0 software.

In this work the numbers of hidden layers was estimated by three different methods. However the best results weren't found by none of these methods, but by a smaller number of neurons than the indicated methods. These results show the independency of theories and postulates in the ANN area. There isn't a simple way to preview the network performance nor how to choose the optimum and unique form its algorithms and functions.

The results showed that the proposed networks are capable of predicting VLE within a reasonable error margin, comparable to methods using specific thermodynamic models, without the necessity to define a model, to calculate interaction parameters or adjustable constants. This result is relevant, respecting the systems presented here have a highly non-ideal behaviour.

## Lista de Figuras

Figura 3.1. Modelos usados para o cálculo do coeficiente de atividade. ....	08
Figura 3.2. Representação da proposta 1. ....	12
Figura 3.3. Representação da proposta 2. ....	12
Figura 3.4. Representação da simplificação da proposta 1.....	13
Figura 3.5. Gráfico dos coeficientes de atividade pela composição do binário etanol (1) -água (2) à pressão de 760 mmHg. ....	13
Figura 4.1. Entrada, junção somadora, função de ativação e saída (Haykin, 1994).....	20
Figura 4.2. Função de transferência logaritmo-sigmóide (logsig).....	21
Figura 4.3. Função de transferência tangente hiperbólica-sigmóide (tansig).....	22
Figura 4.4. Função de transferência linear (purelin). ....	23
Figura 4.5. Exemplo de arquiteturas de rede, a) rede <i>feedforward</i> de uma única camada, b) rede <i>feedforward</i> de múltiplas camadas e c) rede recorrente. ....	24
Figura 6.1. Representação da rede R1. ....	33
Figura 6.2. Comparação entre valores experimentais do ? com valores calculados pela rede neural R1 no treinamento, 20 neurônios na camada escondida.....	36
Figura 6.3. Comparação entre valores experimentais do ? com valores calculados pela rede neural R1 no teste, 20 neurônios na camada escondida.....	36
Figura 6.4. Representação da rede R2. ....	38
Figura 6.5. Algoritmo da temperatura de bolha ( <i>Bolha T</i> ). ....	40
Figura 6.6. Algoritmo da pressão de bolha ( <i>Bolha P</i> ). ....	41
Figura 6.7. Gráfico do ? pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 28. ....	43
Figura 6.8. Gráfico do SSE pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 28.....	43
Figura 6.9. Comparação entre valores experimentais do ? com valores calculados pela rede neural R2 no treinamento, 28 neurônios na camada escondida.....	44

Figura 6.10. Comparação entre valores experimentais do $\hat{y}$ com valores calculados pela rede neural R2 no teste, 28 neurônios na camada escondida.....	45
Figura 6.11. Gráfico do $\hat{y}$ pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 57.....	48
Figura 6.12. Gráfico do SSE pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 57.....	48
Figura 6.13. Comparação entre valores experimentais do $\hat{y}$ com valores calculados pela rede neural R3 no treinamento, 57 neurônios na camada escondida. ....	50
Figura 6.14. Comparação entre valores experimentais do $\hat{y}$ com valores calculados pela rede neural R3 no teste, 57 neurônios na camada escondida.....	50
Figura 6.15. Comparação entre valores experimentais do $\hat{y}$ com valores calculados pela rede neural R3 no treinamento, 18 neurônios na camada escondida.....	51
Figura 6.16. Comparação entre valores experimentais do $\hat{y}$ com valores calculados pela rede neural R3 no teste, 18 neurônios na camada escondida.....	52
Figura 6.17. Proposta 2, Rede R4.....	54
Figura 6.18. Gráfico do $\hat{y}$ pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 63.....	56
Figura 6.19. Gráfico do SSE pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 63.....	56
Figura 6.20. Comparação entre valores experimentais de $y$ e $T$ com valores calculados pela rede neural R4 no treinamento, 63 neurônios na camada escondida. ....	58
Figura 6.21. Comparação entre valores experimentais de $y$ e $T$ com valores calculados pela rede neural R4 no treinamento, 13 neurônios na camada escondida. ....	58
Figura 6.22. Comparação entre valores experimentais de $y$ e $T$ com valores calculados pela rede neural R4 no teste, 63 neurônios na camada escondida. ....	58
Figura 6.23. Comparação entre valores experimentais de $y$ e $T$ com valores calculados pela rede neural R4 no teste, 13 neurônios na camada escondida. ....	58

Figura 6.24. Gráfico do $\sigma$ pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 95.....	60
Figura 6.25. Gráfico do SSE pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 63.....	61
Figura 6.26. Comparação entre valores experimentais de $y$ e $T$ com valores calculados pela rede neural R5 no treinamento, 95 neurônios na camada escondida. ....	62
Figura 6.27. Comparação entre valores experimentais de $y$ e $T$ com valores calculados pela rede neural R5 no treinamento, 19 neurônios na camada escondida. ....	62
Figura 6.28. Comparação entre valores experimentais de $y$ e $T$ com valores calculados pela rede neural R5 no teste, 95 neurônios na camada escondida. ....	63
Figura 6.29. Comparação entre valores experimentais de $y$ e $T$ com valores calculados pela rede neural R5 no teste, 19 neurônios na camada escondida. ....	63

## Lista de Tabelas

Tabela 6.1. Componentes (álcoois-água) e grupos UNIFAC .....	34
Tabela 6.2. Dados de treinamento e teste obtidos pela rede R1 com 20 neurônios na camada escondida.....	36
Tabela 6.3. Exemplo de troca de binários. ....	38
Tabela 6.4. Métodos utilizados para encontrar o número de neurônios na camada escondida para a rede R2.....	42
Tabela 6.5. Dados de treinamento e teste obtidos pela rede R2 com 28 neurônios na camada escondida.....	44
Tabela 6.6. Componentes e grupos UNIFAC usados na rede R3.....	46
Tabela 6.7. Métodos utilizados para encontrar o número de neurônios na camada escondida para a rede R3.....	47
Tabela 6.8. Dados de treinamento e teste obtidos pela rede R3 com 57 neurônios na camada escondida.....	49
Tabela 6.9. Dados de treinamento e teste obtidos pela rede R3 com 18 neurônios na camada escondida.....	51
Tabela 6.10. Métodos utilizados para encontrar o número de neurônios na camada escondida para a rede R4. ....	55
Tabela 6.11. Dados de treinamento e teste da rede R4 com 63 neurônios na camada escondida.....	57
Tabela 6.12. Dados de treinamento e teste da rede R4 com 13 neurônios na camada escondida.....	57
Tabela 6.13. Métodos utilizados para encontrar o número de neurônios na camada escondida para a rede R5. ....	60
Tabela 6.14. Dados de treinamento e teste da rede R5 com 95 neurônios na camada escondida.....	62
Tabela 6.15. Dados de treinamento e teste da rede R5 com 19 neurônios na camada escondida.....	62

# Sumário

<b>1</b>	<b>Introdução</b> .....	<b>1</b>
1.1	Redes neurais artificiais na engenharia química.....	1
1.2	Equilíbrio líquido-vapor x redes neurais artificiais .....	2
1.3	Equilíbrio de fases e propostas de trabalho .....	2
<b>2</b>	<b>Objetivo</b> .....	<b>4</b>
<b>3</b>	<b>Desenvolvimento das propostas de trabalho com os conceitos da termodinâmica</b>	<b>5</b>
3.1	Equilíbrio termodinâmico.....	5
3.2	Coefficiente de atividade .....	6
3.3	Método de contribuição de grupos .....	8
3.4	Método UNIFAC.....	9
3.5	Propostas de trabalho.....	11
3.5.1	Proposta 1 .....	11
3.5.2	Proposta 2 .....	12
3.5.3	Simplificação da proposta 1 .....	13
3.6	Coefficientes de atividade à diluição infinita.....	14
3.7	Revisão bibliográfica: cálculo do ELV através das RNAs.....	15
<b>4</b>	<b>Conceitos de Redes Neurais Artificiais</b> .....	<b>19</b>
4.1	Redes neurais artificiais.....	19
4.2	Modelo de um neurônio artificial .....	19
4.3	Funções de ativação (transferência) .....	21
4.3.1	Função logaritmo-sigmóide (logsig) .....	21
4.3.2	Função tangente hiperbólica-sigmóide (tansig).....	22
4.3.3	Função linear (purelin) .....	22
4.4	Arquiteturas das RNAs.....	23
4.4.1	Rede <i>feedforward</i> com uma única camada.....	23
4.4.2	Rede <i>feedforward</i> de múltiplas camadas .....	24
4.4.3	Redes recorrentes.....	24
4.5	Aprendizado das RNAs .....	24

4.5.1	Aprendizagem supervisionada.....	25
4.5.2	Aprendizagem não-supervisionada.....	25
4.5.3	Aprendizagem por reforço.....	26
4.5.4	Modos de treinamento .....	27
4.6	Treinamento das RNAs .....	27
4.7	Algoritmos de treinamento .....	29
4.7.1	Algoritmo <i>backpropagation</i> .....	29
4.7.2	Algoritmo de Levenberg-Marquardt .....	30
4.7.3	Regularização Bayesiana.....	30
<b>5</b>	<b>Metodologia.....</b>	<b>32</b>
<b>6</b>	<b>Resultados e Discussões.....</b>	<b>33</b>
6.1	Projeto da rede R1 .....	33
6.1.1	Resultados da rede R1 .....	35
6.1.2	Conclusões parciais da rede R1 .....	37
6.2	Projeto da rede R2 .....	37
6.2.1	Temperatura de bolha ( <i>Bolha T</i> ).....	39
6.2.2	Pressão de bolha ( <i>Bolha P</i> ).....	40
6.2.3	Escolha do número de neurônios na camada escondida.....	41
6.2.4	Resultados da rede R2 .....	44
6.2.5	Conclusões parciais da rede R2.....	45
6.3	Projeto da rede R3 .....	46
6.3.1	Escolha do número de neurônios na camada escondida.....	47
6.3.2	Resultados da rede R3 .....	49
6.3.3	Conclusões parciais da rede R3.....	52
6.4	Projeto da rede R4 .....	53
6.4.1	Escolha do número de neurônios na camada escondida.....	54
6.4.2	Resultados da rede R4 .....	57
6.4.3	Conclusões parciais da rede R4.....	59
6.5	Projeto da rede R5 .....	59
6.5.1	Escolha do número de neurônios na camada escondida.....	59
6.5.2	Resultados da rede R5 .....	61

6.5.3	Conclusões parciais da rede R5 .....	63
<b>7</b>	<b>Conclusão .....</b>	<b>65</b>
<b>8</b>	<b>Sugestões para trabalhos futuros .....</b>	<b>67</b>
<b>9</b>	<b>Referências Bibliográficas .....</b>	<b>68</b>

# 1 Introdução

Esse capítulo apresenta a importância do equilíbrio líquido-vapor na engenharia química e as vantagens e desvantagens de se usar as redes neurais artificiais ao invés dos conceitos termodinâmicos no cálculo do equilíbrio líquido-vapor (ELV) de misturas não ideais.

## 1.1 Redes neurais artificiais na engenharia química

A maioria dos processos de engenharia química são não-lineares e complexos e as técnicas de modelagem convencional e simulação dependem frequentemente de certas simplificações no transporte, cinética e termodinâmica. Redes neurais artificiais (RNAs), ao contrário, superam essas limitações da aproximação convencional pela extração da informação desejada diretamente dos dados. Por isso, modelos de processos por RNAs eliminam a necessidade de detalhamento.

Apesar da extensiva aplicação das redes neurais na engenharia química, as RNAs possuem desvantagens, como: tempo excessivo para o treinamento da rede e o *overfitting* (sobre-ajuste), que ocorre quando a rede perde a sua capacidade de generalização.

A habilidade de reconhecer e reproduzir as relações de causa e efeito através do treinamento, e a filtragem dos ruídos ou dados irrelevantes, e ainda a possibilidade de se trabalhar com múltiplos vetores de entrada e saída, fazem da rede neural um eficiente representante dos sistemas não lineares. Os modelos baseados nesse método podem ser desenvolvidos para usar como entradas, variáveis que são familiares e de fácil acesso (pressão e temperatura, por exemplo), em vez de parâmetros que não são facilmente medidos, enquanto que como saídas pode-se usar variáveis de difícil obtenção (coeficientes de atividade, por exemplo).

## 1.2 Equilíbrio líquido-vapor x redes neurais artificiais

O estudo do equilíbrio líquido-vapor é fundamental para o dimensionamento, modelagem e simulação de processos e equipamentos na engenharia química. Dados de ELV são estimados geralmente usando modelos termodinâmicos baseados em fundamentos bem conhecidos como o critério de igualdade de potenciais químicos em ambas as fases. A maioria dos modelos termodinâmicos possui equações empíricas ou semi-empíricas, das quais cada uma tem várias constantes a se determinar e precisam de regras de mistura. As incertezas introduzem um limite nos sistemas e condições para os quais os métodos da termodinâmica podem ser usados seguramente. As equações que se aplicam bem para sistemas de hidrocarbonetos, por exemplo, não apresentam as mesmas vantagens para sistemas contendo componentes polares. As equações também não se aplicam satisfatoriamente para a região crítica nem para as propriedades líquidas. No entanto, cada uma dessas limitações tem aplicabilidade em diferentes sistemas, por exemplo, enquanto o método UNIFAC é aplicável para sistemas contendo água, este tem suas limitações com respeito a misturas de hidrocarbonetos. A dificuldade na análise das misturas cresce com a não idealidade.

As RNAs por serem de natureza puramente numérica não requerem modelos termodinâmicos e, portanto, são convenientes para a predição de dados de ELV. Além disso, um banco de dados limitado pode ser usado para treinar a rede adequadamente e para testá-la (Sharma *et al.*, 1999).

## 1.3 Equilíbrio de fases e propostas de trabalho

Uma grande parte do dimensionamento de equipamentos na engenharia química é relativo a processos de separação. Muitos desses são operações difusionais que envolvem equilíbrio de fases, das quais destilação, absorção e extração líquido-líquido são as mais comuns. Para um dimensionamento adequado desses processos de separação, precisa-se de informações quantitativas do equilíbrio em misturas binárias ou multicomponentes. (Fredenslund *et al.*, 1977).

Esse trabalho tem como propósito o desenvolvimento de um método para prever o ELV de um modo diferente, através das redes neurais artificiais. Ao longo desse trabalho serão descritos os objetivos global e específicos no Capítulo 2 e a termodinâmica do ELV em conjunto com as propostas de trabalho dessa dissertação no Capítulo 3. Os conceitos e as definições das RNAs serão descritos no Capítulo 4 e a metodologia no Capítulo 5. Os resultados e análises das redes serão abordados no Capítulo 6, a conclusão do trabalho no Capítulo 7 e as sugestões para trabalhos futuros no Capítulo 8.

## 2 Objetivo

Essa dissertação tem como objetivo global o estudo da viabilidade de se utilizar redes neurais artificiais (RNAs) para a predição do equilíbrio líquido-vapor (ELV) através da abordagem do método de contribuição de grupos, no caso o método UNIFAC por este ser amplamente utilizado.

Como objetivos específicos foram sugeridos:

- Uma primeira proposta de trabalho (proposta 1) que visa à substituição do método UNIFAC para o cálculo do coeficiente de atividade pelas redes neurais artificiais;
- Uma segunda proposta de trabalho (proposta 2) que calcula diretamente a composição da fase vapor através das redes neurais artificiais;
- Comparar o número de neurônios na camada escondida obtidos por três métodos diferentes.

As duas propostas de trabalho mencionadas acima serão primeiro implementadas para misturas binárias álcoois-água com cinco grupos funcionais UNIFAC. Após isso, verificada a capacidade das RNAs calcularem as variáveis propostas, os dados serão estendidos para diversos componentes (álcoois, água, aldeídos e cetonas) e mais três grupos funcionais UNIFAC.

### 3 Desenvolvimento das propostas de trabalho com os conceitos da termodinâmica

Esse capítulo apresenta a revisão bibliográfica que serviu como base para esse trabalho, os tópicos da termodinâmica serão abordados juntamente com as propostas de trabalho sugeridas nessa dissertação. Esse capítulo também apresenta uma revisão de artigos publicados na área de RNAs e ELV.

#### 3.1 Equilíbrio termodinâmico

Para todo o componente  $i$  na mistura, a condição de equilíbrio termodinâmico é dada pela Equação (3.1):

$$f_i^V = f_i^L \quad (3.1)$$

Na equação acima,  $f$  é a fugacidade,  $V$  é a fase vapor e  $L$  é a fase líquida.

Através de simplificações chega-se à equação de Raoult estendida, Eq. (3.2):

$$y_i P = \gamma_i x_i P_i^{sat} \quad i = 1, 2, \dots, n \quad (3.2)$$

Nas equações acima,  $y_i$  é a composição de  $i$  na fase vapor,  $P$  é a pressão do sistema,  $\gamma_i$  é o coeficiente de atividade do componente  $i$ ,  $x_i$  é a composição de  $i$  na fase líquida e  $P_i^{sat}$  é a pressão de saturação do componente  $i$ . O desenvolvimento e simplificações das equações de equilíbrio de fases podem ser encontrados em referências como: Sandler (1989), Smith *et al.* (2000), Reid *et al.* (1987), Fredenslund *et al.* (1977) e Walas (1985).

A equação de Raoult estendida é usada para o cálculo do ELV, sendo que a variável mais difícil de se obter é o  $\gamma$ , pois  $P$  é facilmente medido e  $P^{\text{sat}}$  pode ser calculado pela equação de Antoine para uma dada temperatura.

## 3.2 Coeficiente de atividade

O coeficiente de atividade ( $\gamma$ ) representa a não idealidade da mistura na fase líquida, sendo que, para misturas ideais,  $\gamma = 1$ . O  $\gamma$  é relacionado com a energia de Gibbs de excesso ( $G^E$ ), através das equações:

$$G^E = RT(n_1 \ln \gamma_1 + n_2 \ln \gamma_2) \quad (3.3)$$

$$RT \ln \gamma_1 = \left( \frac{\partial G^E}{\partial n_1} \right)_{T,P,n_2} \quad (3.4)$$

$$RT \ln \gamma_2 = \left( \frac{\partial G^E}{\partial n_2} \right)_{T,P,n_1} \quad (3.5)$$

As funções de excesso fornecem um método termodinâmico consistente para dados experimentais de mistura binária e também para misturas multicomponentes. Frequentemente, no entanto, poucos ou nenhum dado de mistura são obtidos, e é necessário estimar os coeficientes de atividade por algumas correlações (Reid *et al.*, 1987).

Essas equações foram propostas para correlacionar os coeficientes de atividade com a composição e em uma menor extensão com a temperatura, como, por exemplo, as equações de Margules, van Laar, Wilson, NRTL e UNIQUAC (essas equações podem ser encontradas em Reid *et al.* (1987)). Todas essas expressões contêm constantes ajustáveis das quais, pelo menos a princípio, dependem da temperatura. Essa dependência pode ser em alguns casos desprezada, especialmente se o intervalo de temperatura for pequeno. Na prática, o número de constantes ajustáveis por binário são normalmente duas ou três. Quanto maior o número de constantes, melhor a representação dos dados, mas é preciso um número maior de pontos experimentais confiáveis para determiná-las.

Para muitas misturas binárias moderadamente não ideais, praticamente todas as correlações que contém dois (ou mais) parâmetros binários dão bons resultados. A única razão de se escolher uma ou outra equação é do ponto de vista matemático, já que as mais antigas (Margules, van Laar) são matematicamente mais fáceis de serem implementadas do que as mais novas (Wilson, NRTL e UNIQUAC). A equação de Margules de dois sufixos (um parâmetro) é aplicável somente para misturas simples onde os componentes são de natureza química e tamanho molecular parecidos (Reid *et al.*, 1987).

Para misturas altamente não ideais, por exemplo, soluções de álcoois com hidrocarbonetos, a equação de Wilson é provavelmente a mais utilizada (Reid *et al.*, 1987), pois, ao contrário da equação NRTL, Wilson contém somente dois parâmetros ajustáveis e é matematicamente mais simples do que a equação UNIQUAC. Para estas misturas, a equação de Margules com três sufixos e a equação de van Laar representam os dados com pouco sucesso. A equação de Margules com quatro sufixos (três parâmetros) não apresenta vantagem significativa em relação à equação NRTL.

Ao contrário da equação de Wilson original, as equações NRTL e UNIQUAC são aplicáveis tanto para o equilíbrio líquido-vapor quanto para o equilíbrio líquido-líquido. Apesar da equação UNIQUAC ser matematicamente mais complexa do que a NRTL, ela apresenta a vantagem de possuir somente dois (ao invés de três) parâmetros ajustáveis. Além disso, a equação UNIQUAC é aplicável para soluções contendo pequenas ou grandes moléculas, incluindo polímeros.

Segundo Fredenslund *et al.* (1977), os modelos citados acima têm três pontos em comum:

1. O equilíbrio líquido-vapor pode ser predito usando informações experimentais somente de sistemas binários;
2. Eles permitem extrapolação com respeito à temperatura, pelo menos dentro de uma faixa limitada de aplicação;
3. Seu uso depende da disponibilidade dos parâmetros binários calculados dos dados de equilíbrio de fases experimental para todas as possíveis combinações binárias na mistura multicomponente.

Outra forma de se calcular o coeficiente de atividade foi proposto por Fredenslund *et al.* (1977). O método UNIFAC é baseado na equação UNIQUAC e no método de contribuição de grupos.

A Figura 3.1 mostra os principais modelos para o cálculo do  $\gamma$ .

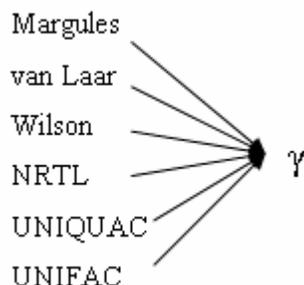


Figura 3.1 - Modelos usados para cálculo do coeficiente de atividade.

### 3.3 Método de contribuição de grupos

Para o cálculo das propriedades termodinâmicas é conveniente considerar uma molécula como um grupo funcional agregado. Como resultado, algumas propriedades termodinâmicas de fluidos puros, por exemplo, capacidade calorífica e volume crítico, podem ser calculados pela soma das contribuições de cada grupo. A extensão desse conceito para misturas foi sugerida, e várias tentativas foram feitas para estabelecer o método de contribuição de grupos para calores de mistura e coeficientes de atividade (Reid *et al.*, 1987).

A idéia básica do método de contribuição de grupos é que, considerando que existem milhares de componentes químicos de interesse na tecnologia química, o número de grupos funcionais que constituem esses componentes é muito pequeno. Portanto, se for assumido que uma propriedade física de um fluido é a soma das contribuições feitas pelos grupos funcionais da sua molécula, ter-se-á obtido uma técnica possível para correlacionar as propriedades de um número grande de fluidos em termos de um número pequeno de parâmetros que caracterizam as contribuições dos grupos individuais. Além disso, obter-se-ia uma técnica que prediz o equilíbrio de fases em um sistema dos quais não existem dados (Fredenslund *et al.*, 1977).

Qualquer método de contribuição de grupos é aproximado, porque a contribuição dada por um grupo em uma molécula não é necessariamente a mesma em outra molécula. O pressuposto fundamental do método de contribuição de grupos é a aditividade, a contribuição feita por um grupo é assumida como independente da que é feita em outro grupo. Este pressuposto é válido somente quando a influência de um grupo na molécula não é afetada pela natureza de outros grupos dentro da molécula. Por exemplo, não se espera que a contribuição do grupo C=O (carbonil) na acetona seja a mesma que a contribuição do C=O no ácido acético. De outro modo, a experiência sugere que a contribuição do grupo C=O na acetona seja aproximadamente a mesma que a contribuição do C=O na 2-butanona (Fredenslund *et al.*, 1977).

A precisão da correlação melhora com o crescimento dos grupos. No limite, onde mais e mais distinções são feitas, volta-se à molécula original. Com isso, a vantagem do método de contribuição de grupos é perdida. Então o número de grupos distintos deve permanecer pequeno, mas não tão pequeno que os efeitos das propriedades físicas da estrutura molecular sejam desprezados (Fredenslund *et al.*, 1977).

A extensão da idéia do método de contribuição de grupos para misturas é extremamente atrativa, já que o número de substâncias puras na tecnologia química é muito grande, o número de diferentes misturas é muito maior, em vários graus de magnitude (Fredenslund *et al.*, 1977). Assim, através do método de contribuição de grupos as propriedades de um grande número de misturas multicomponentes de interesse na indústria química podem ser constituídas de, por exemplo, cinco grupos funcionais (Walas, 1985).

### **3.4 Método UNIFAC**

O método UNIFAC trabalha com o conceito de que os coeficientes de atividade em misturas são relacionados com interações entre grupos estruturais. A vantagem do método UNIFAC é que ele é capaz de prever através de alguns parâmetros os coeficientes de atividade de sistemas que não foram estudados experimentalmente, mas que contém o mesmo grupo funcional.

Tanto no método UNIFAC quanto no método UNIQUAC trabalha-se com o coeficiente de atividade molecular, que é dividido em duas partes: uma parte combinatorial e a outra residual, Equação (3.6).

No modelo UNIQUAC a parte combinatorial representa as diferenças em tamanho e forma das moléculas na mistura e a parte residual representa as energias de interação, o tamanho do grupo funcional e a área de superfície de interação, que são introduzidas de uma obtenção independente, dados estruturais moleculares de componentes puros (Reid *et al.*, 1987).

A equação UNIQUAC normalmente apresenta uma boa representação de ambos, equilíbrio líquido-vapor e equilíbrio líquido-líquido de misturas binárias e multicomponentes, como: hidrocarbonetos, cetonas, ésteres, água, aminas, álcoois, nitrilas, etc (Reid *et al.*, 1987). Numa mistura multicomponente, a equação UNIQUAC para o coeficiente de atividade do componente  $i$  (molecular) é:

$$\ln g_i = \ln g_i^C + \ln g_i^R \quad (3.6)$$

$$\ln g_i^C = \ln \frac{f_i}{x_i} + \frac{z}{2} q_i \ln \frac{q_i}{f_i} + \ell_i \frac{f_i}{x_i} \sum_j x_j \ell_j \quad (3.7)$$

$$\ln g_i^R = q_i \left[ 1 - \ln \left( \sum_j q_j t_{ji} \right) - \sum_j \frac{q_j t_{ij}}{\sum_k q_k t_{kj}} \right] \quad (3.8)$$

Nessas equações  $x_i$  é a fração molar do componente  $i$ ,  $q_i$  é a fração de área, e  $F_i$  é a fração de segmento, que é similar à fração volumétrica. Parâmetros de componentes puros  $r_i$  e  $q_i$ , são, respectivamente, medidas de volume molecular de van der Waals e áreas de superfície molecular. Na equação UNIQUAC, os dois parâmetros binários ajustáveis,  $t_{ij}$  e  $t_{ji}$ , que aparecem na Equação (3.8), são retirados de dados de equilíbrio experimentais.

No método UNIFAC, a parte combinatorial dos coeficientes de atividade da equação UNIQUAC, Equação (3.7), é usado diretamente. Somente propriedades dos componentes puros entram nesta equação (Reid *et al.*, 1987).

A parte residual do coeficiente de atividade, Equação (3.8), é substituída pelo conceito de contribuição de grupos. Em vez da Equação (3.8), escreve-se:

$$\ln g_i^R = \sum_k n_k^{(i)} (\ln \Gamma_k - \ln \Gamma_k^{(i)}) \quad (3.9)$$

Onde  $\Gamma_k$  é o coeficiente de atividade do grupo residual e  $\Gamma_k^{(i)}$  é o coeficiente de atividade residual do grupo k em referência à solução contendo somente moléculas do tipo i.

O método UNIFAC pode ser aplicado para misturas binárias ou multicomponentes onde o modelo UNIQUAC se aplica, ou seja, fora da região crítica. Além disso, todos os componentes têm que ser condensáveis. Finalmente, o método UNIFAC não se aplica a misturas contendo polímeros e é raramente aplicado para misturas contendo componentes que usem mais de dez grupos funcionais (Fredenslund *et al.*, 1977).

### 3.5 Propostas de trabalho

Nessa seção serão revisadas as propostas de trabalho já mencionadas no objetivo dessa dissertação, Capítulo 2.

#### 3.5.1 Proposta 1

Como visto anteriormente, a primeira proposta de trabalho dessa dissertação visa o cálculo do coeficiente de atividade (?) através da substituição do modelo UNIFAC pelas RNAs, Figura 3.2.

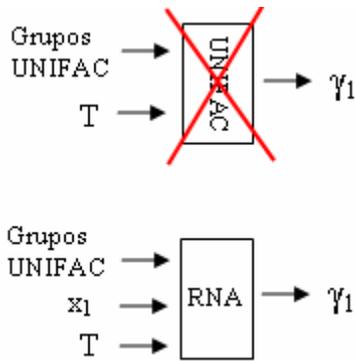


Figura 3.2 - Representação da proposta 1.

Na figura acima,  $x_1$  é a composição na fase líquida do componente 1 e  $T$  é a temperatura do sistema em kelvin.

### 3.5.2 Proposta 2

A segunda proposta dessa dissertação tem como objetivo calcular a composição na fase vapor ( $y$ ) diretamente através das RNAs, Figura 3.3.

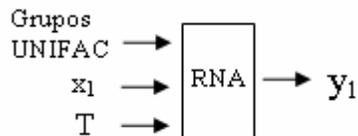


Figura 3.3 - Representação da proposta 2.

Essas propostas foram aplicadas primeiramente para misturas binárias álcoois-água com cinco grupos funcionais UNIFAC e depois foram estendidas para outros componentes (álcoois, água, aldeídos e cetonas) com mais três grupos funcionais UNIFAC. As duas propostas calculam as suas variáveis para dados isobáricos e isotérmicos que estão na faixa de 10,00 a 215,47 °C e 0,739 a 7109,54 mmHg.

### 3.5.3 Simplificação da proposta 1

Analisando-se a proposta 1, percebeu-se que seria possível fazer uma simplificação na mesma se fosse usado, ao invés do coeficiente de atividade, o coeficiente de atividade à diluição infinita ( $\gamma^\infty$ ). O  $\gamma^\infty$  representa os extremos das composições, quando a composição tende a 0 ou a 1, assim, usando essa variável ter-se-ia uma entrada a menos na rede ( $x_1$ ). A Figura 3.4 mostra a simplificação da proposta 1. A Figura 3.5 mostra o ELV da mistura binária etanol-água à 50°C, a figura indica os coeficientes de atividade à diluição infinita (nos extremos, quando  $x \rightarrow 0$  e  $x \rightarrow 1$ ) e os coeficientes de atividade que dependem da composição da mistura.

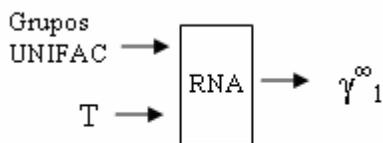


Figura 3.4 - Representação da simplificação da proposta 1.

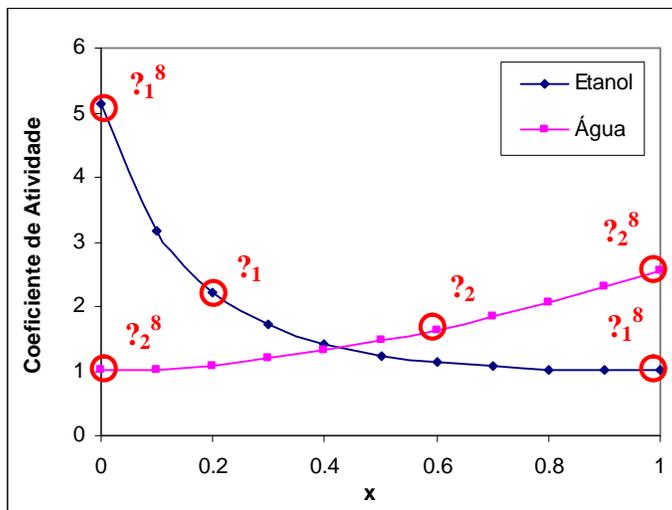


Figura 3.5 - Gráfico dos coeficientes de atividade pela composição do binário etanol (1) – água (2) à pressão de 760 mmHg.

Através do  $\gamma^\infty$  pode-se calcular o  $\gamma$  para qualquer composição e, usando a Eq. (3.2), o ELV da mistura.

### 3.6 Coeficientes de atividade à diluição infinita

Os coeficientes de atividade à diluição infinita ( $\gamma^{\infty}$ ) representam os coeficientes de atividade ( $\gamma$ ) nos extremos das composições, quando  $x_1 \rightarrow 0$  e  $x_2 \rightarrow 1$ , vide Figura 3.5. Coeficientes de atividade à diluição infinita experimentais são particularmente usados para calcular parâmetros das expressões da energia de Gibbs de excesso. Segundo Reid *et al.* (1987), eles podem ser usados para avaliar duas constantes ajustáveis em qualquer expressão desejada para  $g^E$ . Por exemplo, considerando a equação de van Laar (Eq. 3.11):

$$g^E = Ax_1x_2 \left( x_1 \frac{A}{B} + x_2 \right)^{-1} \quad (3.11)$$

Através das Equações (3.3), (3.4) e (3.5) chega-se às equações:

$$RT \ln g_1 = A \left( 1 + \frac{A}{B} \frac{x_1}{x_2} \right)^{-2} \quad (3.12)$$

$$RT \ln g_2 = A \left( 1 + \frac{A}{B} \frac{x_2}{x_1} \right)^{-2} \quad (3.13)$$

No limite, onde  $x_1 \rightarrow 0$  ou  $x_2 \rightarrow 0$ , as Equações (3.12) e (3.13) se tornam:

$$RT \ln g_1^{\infty} = A \quad (3.14)$$

$$RT \ln g_2^{\infty} = B \quad (3.15)$$

Calcular os parâmetros da equação através de  $\gamma^{\infty}$  é particularmente simples para a equação de van Laar, mas, à princípio, cálculos parecidos podem ser feitos usando as equações de dois parâmetros para a energia de Gibbs de excesso. Se a equação de três

parâmetros, isto é, a NRTL, é usada, um método independente pode ser escolhido para determinar o terceiro parâmetro  $\alpha_{12}$  (Reid *et al.*, 1987).

### 3.7 Revisão bibliográfica: cálculo do ELV através das RNAs

Vários pesquisadores estudaram a aplicação das RNAs no cálculo de propriedades termodinâmicas de misturas, como, por exemplo, no ELV. Nessa revisão serão apontados alguns artigos publicados nessa área importantes para o desenvolvimento deste trabalho.

Petersen *et al.* (1994) desenvolveram um novo método de contribuição de grupos para o cálculo do coeficiente de atividade (?) utilizando RNAs. Foram construídas duas redes, uma com uma camada escondida com quatro neurônios e outra somente com a camada de entrada e a camada de saída. Os resultados das redes foram comparados com o método UNIFAC e dados experimentais, o método UNIFAC se mostrou superior ao método utilizando RNAs.

Sharma *et al.* (1999) utilizaram RNAs na predição do equilíbrio líquido-vapor (ELV) em dois sistemas binários, metano-etano e amônia-água. Foram usados como dados de entrada na rede a pressão ( $P$ ) e a temperatura ( $T$ ) e como saída a composição da fase líquida ( $x$ ) e a composição na fase vapor ( $y$ ). Foi feita uma rede para cada sistema, as redes possuem uma camada escondida com treze neurônios tanto para o sistema amônia-água quanto para o sistema metano-etano, foi usado o algoritmo *backpropagation* como função de treinamento. A rede foi treinada também com as entradas  $P$  e  $T$  e saída  $x$ , e com entradas  $P$  e  $T$  e saída  $y$ , essas últimas redes apresentaram melhores resultados do que as primeiras, os erros nos dois sistemas estudados foram de  $\pm 1\%$ .

Iliuta *et al.* (2000) utilizaram duas RNAs para prever o ELV de sistemas com dois solventes e um eletrólito, uma rede com saída  $y$  e  $P$  com dezesseis neurônios na camada escondida e outra rede com  $y$  e  $T$  e dezoito neurônios na camada escondida, como função de ativação usou-se uma função sigmóide. Os resultados foram comparados com valores experimentais e foram considerados bons pelos autores. As RNAs se mostraram uma boa ferramenta para a predição de ELV de misturas altamente não ideais.

Piotrowski *et al.* (2003) aplicaram RNAs para prever o ELV em um sistema industrial complexo, o processo de síntese de uréia. Foram usados 11 neurônios na camada escondida, três neurônios de entrada e quatro neurônios de saída, entre eles o  $y$  da  $\text{NH}_3$ , do  $\text{CO}_2$  e da  $\text{H}_2\text{O}$ . Usou-se o algoritmo *backpropagation* como algoritmo de treinamento e a função sigmóide como função de ativação. As RNAs obtiveram um resultado satisfatório para o uso nos cálculos de engenharia.

Urata *et al.* (2002) aplicaram RNAs para estimar o ELV em sistemas binários contendo hidrofluoréters. O cálculo foi feito em duas etapas, a primeira etapa foi constituída de uma rede para classificar o  $\ln(?)$  em positivo e negativo. A segunda etapa foi constituída de duas redes, uma que prevê o  $\ln(?_1)$  para o caso positivo e outra para o caso negativo. Todas as redes tanto para a primeira quanto para a segunda etapa utilizaram uma camada escondida, algoritmo de treinamento *backpropagation* e função de ativação sigmóide. Os resultados das redes foram considerados bons, com erros pequenos. Mas a rede apresentou um problema, ela não tem como garantir que o  $?_1$  seja igual a 1 quando  $x_1$  é igual a 1. Isso acontece porque as RNAs não seguem nenhum modelo termodinâmico e não se podem restringir os seus valores de saída.

Ganguly (2003) utilizou uma RNA do tipo função de base radial (FBR) para a predição do ELV de quatro sistemas binários e dois sistemas ternários. Foram propostas duas configurações de rede, a primeira utiliza somente uma RNA e possui como entradas  $x$  e  $P$  ou  $T$  e como saídas  $y$  e  $P$  ou  $T$ ; a segunda configuração é composta de duas redes neurais em série, a primeira tem como entrada de rede as variáveis  $x$  e  $P$  ou  $T$ , como saída da primeira rede e entrada da segunda as variáveis  $P$ ,  $T$ ,  $x$  e como saída  $y$ . Os resultados das RNAs foram comparados com dados UNIQUAC retirados da Coleção DECHEMA, a precisão dos dois métodos foi considerada similar. Foi comparado também o número de parâmetros utilizado pelas RNAs com o utilizado pela equação UNIQUAC, sendo que este último utiliza mais parâmetros do que o primeiro, o que demonstra uma certa vantagem de se usar RNAs na predição do ELV. Chegou-se também à conclusão que a configuração com duas RNAs em série não apresentou nenhuma vantagem sobre a configuração com uma única RNA.

Mohanty (2006) estimou o ELV do sistema binário dióxido de carbonodifluórometano, normalmente usados como refrigerantes. Foi construída uma rede com uma

camada escondida com a configuração 2-10-2, utilizando como variáveis de entrada  $T$  e  $x_{\text{CO}_2}$  e como saída da rede  $P$  e  $y_{\text{CO}_2}$ . O autor comparou os resultados das RNAs com valores obtidos da equação de estado de Peng-Robinson, os valores obtidos pelas redes neurais foram considerados bons, mas com as desvantagens de que a RNA é limitada ao sistema e faixa de dados estudado e também não pode ser usada para estimar outras propriedades termodinâmicas além daquela inicialmente estabelecida, diferentemente das equações de estado. Outra desvantagem é que as RNAs precisam de muitos dados experimentais de qualidade para o treinamento e teste da rede.

Mohanty (2005) empregou as RNAs em três sistemas binários, dióxido de carbonocaproato de etila, caprilato de etila e caprato de etila os quais são muito utilizados em extração supercrítica. Foram usadas RNAs com oito neurônios na camada escondida para o sistema dióxido de carbonocaproato de etila, cinco para o sistema dióxido de carbonocaprilato de etila e seis para dióxido de carbonocaprato de etila. Foram testadas as funções de ativação “tansig” e “logsig” para a camada escondida e linear para a camada de saída, o algoritmo de treinamento usado foi o *backpropagation*. A rede apresentou bons resultados, com erros menores do que os apresentados pelas equações de Peng-Robinson e Soave-Redlich-Kwong, a função “tansig” obteve melhores resultados do que a “logsig”.

Nguyen *et al.* (2007) empregaram RNAs na predição do ELV de sistemas ternários sendo um dos componentes um sal. Os autores trabalharam com diversas redes, com uma e duas camadas escondidas e funções de ativação “logsig”, “tansig” e “purelin” (linear). A melhor RNA encontrada foi a rede com a configuração 8-7-6-4 (número de neurônios na camada de entrada, primeira camada escondida, segunda camada escondida e camada de saída), as funções de ativação usadas foram logsig-logsig-purelin. A RNA obteve resultados tão bons quanto os do modelo teórico termodinâmico (Tan-Wilson), mas usando um número de parâmetros muito menor, a rede também obteve resultados considerados bons quando comparados com dados experimentais.

Analisando os trabalhos publicados percebeu-se que foram poucos os artigos que estudaram a predição do ELV através das RNAs para vários sistemas, ou seja, vários grupos de componentes como, por exemplo, cetonas, aldeídos, álcoois. A maioria dos trabalhos estudou sistemas binários ou ternários simples com apenas um grupo de componentes. Mesmo os trabalhos que usaram mais grupos, geralmente eram grupos

similares ou que se prestavam a um mesmo fim, como por exemplo, em alguma reação específica.

Nessa dissertação foi proposto um trabalho diferente, em que vários grupos são preditos em uma mesma rede, tanto para dados isobáricos quanto para dados isotérmicos, tentando-se assim fazer uma rede mais complexa mas com um uso mais geral.

Percebeu-se também que com exceção do artigo de Petersen *et al.* (1993) que desenvolveram uma nova metodologia de cálculo do coeficiente de atividade utilizando RNAs, os artigos se limitaram a calcular o ELV diretamente (obtendo  $x$  e/ou  $y$ ) para misturas binárias ou ternárias. Nessa dissertação propôs-se além do cálculo direto do  $y$ , o cálculo do  $\gamma$ , obtendo-se assim o ELV de uma forma indireta.

## **4 Conceitos de Redes Neurais Artificiais**

Esse capítulo apresenta a revisão da literatura de redes neurais artificiais, do tipo perceptron de múltiplas camadas, que serviu como base na presente dissertação.

### **4.1 Redes neurais artificiais**

Uma rede neural artificial é um circuito composto por uma grande quantidade de unidades simples, chamados neurônios artificiais, sendo que cada unidade apresenta um comportamento específico de entrada-saída, determinado pela sua função de transferência, pelas interconexões com outras unidades, e possivelmente pelas entradas externas.

As aplicações das redes neurais artificiais são as mais diversas e envolvem problema de otimização, agrupamento de dados, classificação, regressão e reconhecimento de padrões. Devido ao grande número de problemas que as RNAs são capazes de resolver pode-se aplicá-las nas mais diversas áreas, não somente na engenharia, mas também na medicina, economia, finanças, entretenimento e outras.

### **4.2 Modelo de um neurônio artificial**

O elemento computacional básico empregado na maioria das redes neurais é um integrador. Trata-se de um elemento processador de informações que é fundamental para a operação das redes neurais.

As principais partes de um neurônio artificial são:

- as sinapses, caracterizadas pelos pesos associados;
- a junção somadora;
- a função de ativação.

A Figura 4.2 é a representação do integrador:

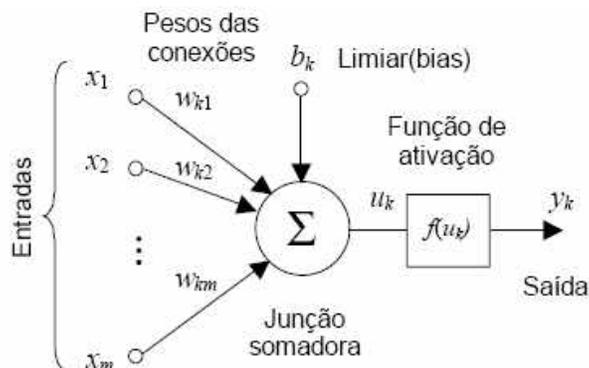


Figura 4.1 - Entrada, junção somadora, função de ativação e saída (Haykin, 1994).

Nessa representação, identificam-se três elementos básicos do modelo neuronal:

1. Um conjunto de sinapses, cada uma caracterizada por um peso. Especificamente, um sinal  $x_j$  na entrada da sinapse  $j$  conectada ao neurônio  $k$  é multiplicado pelo peso sináptico  $w_{kj}$ . É importante notar a maneira como são escritos os índices dos pesos sinápticos  $w_{kj}$ . O primeiro índice se refere ao neurônio em questão e o segundo se refere ao terminal de entrada da sinapse à qual o peso se refere. O peso sináptico de um neurônio artificial pode estar em um intervalo que inclui valores negativos bem como positivos.
2. Um somador para somar os sinais de entrada, ponderados pelas respectivas sinapses do neurônio, as operações descritas aqui constituem um combinador linear.
3. Uma função de ativação é usada para restringir a amplitude da saída de um neurônio. A função de ativação é também dita como função restritiva já que restringe (limita) o intervalo permissível de amplitude do sinal de saída a um valor finito. Normalmente o intervalo normalizado da amplitude da saída de um neurônio é escrito como o intervalo unitário fechado  $[0,1]$  ou  $[-1,1]$ .

O modelo neuronal da Figura 4.2 inclui também um bias aplicado externamente, representado por  $b_k$ . O bias tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se ele é positivo ou negativo, respectivamente.

Em termos matemáticos, pode-se descrever um neurônio  $k$  escrevendo o seguinte par de equações:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (4.5)$$

$$y_k = f(u_k + b_k) \quad (4.6)$$

Nas equações acima,  $x_1, x_2, \dots, x_m$  são sinais de entrada;  $w_{k1}, w_{k2}, \dots, w_{km}$  são os pesos sinápticos do neurônio  $k$ ;  $u_k$  é a saída do combinador linear devido aos sinais de entrada;  $b_k$  é o bias;  $f(\cdot)$  é a função de ativação; e  $y_k$  é o sinal de saída do neurônio.

### 4.3 Funções de ativação (transferência)

A função de ativação ou de transferência da Figura 4.2 pode ser uma função linear ou não linear de  $u_k$ . Existe uma variedade de funções de transferência, três das mais utilizadas são mostradas a seguir.

#### 4.3.1 Função logaritmo-sigmóide (logsig)

A Equação (4.7) e a Figura 4.3 representam a função de ativação logaritmo-sigmóide (logsig).

$$y_k = \frac{1}{1 + e^{-u_k}} \quad (4.7)$$

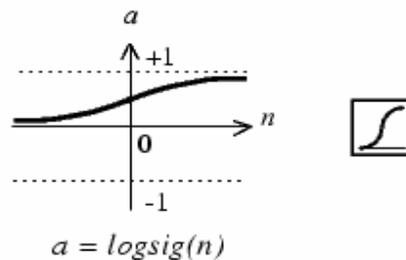


Figura 4.2 - Função de transferência logaritmo-sigmóide (logsig) (Hagan, 1996).

### 4.3.2 Função tangente hiperbólica-sigmóide (tansig)

As Equações (4.8) e (4.9) e a Figura 4.4 representam a função de ativação tangente hiperbólica-sigmóide (tansig). A Equação (4.8) foi retirada da literatura e a Equação (4.9) foi retirada do Matlab, essas equações são equivalentes.

$$y_k = \frac{e^{u_k} - e^{-u_k}}{e^{u_k} + e^{-u_k}} \quad (4.8)$$

ou

$$y_k = \left( \frac{2}{1 + e^{-2u_k}} \right) - 1 \quad (4.9)$$

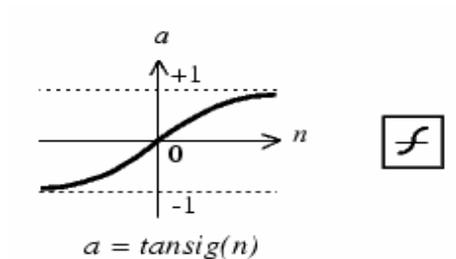


Figura 4.3 - Função de transferência tangente hiperbólica-sigmóide (tansig) (Hagan, 1996).

### 4.3.3 Função linear (purelin)

A Equação (4.10) e a Figura 4.5 representam a função de ativação linear (purelin). A saída da função de transferência linear é igual a sua entrada.

$$y_k = u_k \quad (4.10)$$

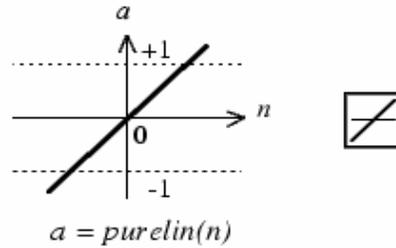


Figura 4.4 - Função de transferência linear (purelin) (Hagan, 1996).

## 4.4 Arquiteturas das RNAs

A maioria das RNAs utiliza arquiteturas padronizadas, projetadas especialmente para resolver algumas classes de problemas.

Existem basicamente três camadas em uma rede neural artificial do tipo perceptron: uma camada de entrada, uma camada intermediária, e uma camada de saída. Entretanto, nem todas as RNAs possuem camadas intermediárias.

A forma como os neurônios estão interconectados está intimamente relacionada ao algoritmo a ser utilizado no seu treinamento. Existem, basicamente, três tipos principais de arquiteturas em RNAs: redes *feedforward* de uma única camada, redes *feedforward* de múltiplas camadas e redes recorrentes.

### 4.4.1 Rede *feedforward* com uma única camada

Esse caso mais simples de rede em camadas consiste em uma camada de entrada e uma camada de saída. Geralmente os neurônios de entrada são lineares, ou seja, eles simplesmente propagam o sinal de entrada para a próxima camada. São também denominados de neurônios sensoriais.

Essa rede é denominada *feedforward* porque a propagação do sinal ocorre apenas da entrada para a saída, ou seja, é apenas no sentido positivo.

#### 4.4.2 Rede *feedforward* de múltiplas camadas

As redes de múltiplas camadas possuem uma ou mais camadas intermediárias ou escondidas. A saída de cada camada intermediária é utilizada como entrada para a próxima camada. Em geral, o algoritmo de treinamento para esse tipo de rede envolve a retropropagação do erro entre a saída da rede e uma saída desejada conhecida.

#### 4.4.3 Redes recorrentes

Esse tipo de rede se difere das demais pois é a única que permite realimentação, que pode ocorrer entre unidades e/ ou entre camadas.

A Figura 4.1 mostra as três redes descritas acima:

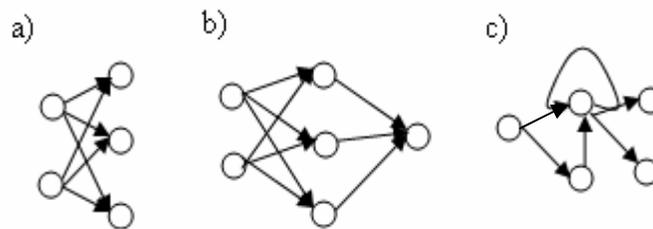


Figura 4.5 - Exemplo de arquiteturas de rede, a) rede *feedforward* de uma única camada, b) rede *feedforward* de múltiplas camadas e c) rede recorrente.

### 4.5 Aprendizado das RNAs

A propriedade mais importante das redes neurais é a habilidade de aprender de seu ambiente e com isso melhorar seu desempenho. Isso é feito através de um processo iterativo de ajustes aplicado a seus pesos, o treinamento. O aprendizado ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas.

Uma das características das RNAs é o local onde o conhecimento está armazenado. Nos casos mais simples, esse conhecimento é armazenado nos pesos das conexões entre neurônios.

A representação de conhecimento é feita de forma que o conhecimento necessariamente influencie a forma de processamento da rede, ou seja, o seu comportamento de entrada-saída. Se o conhecimento está armazenado nos pesos das conexões, então o processo de aprendizagem corresponde a identificar um conjunto apropriado de pesos de forma que a rede se comporte como o desejado. Esse processo pode acontecer de diversas formas, dependendo do paradigma, do algoritmo e do modo utilizado. A seguir serão mostrados os três paradigmas de aprendizado mais utilizados e os modos de treinamento.

#### **4.5.1 Aprendizagem supervisionada**

Esse processo envolve a modificação dos valores dos pesos da rede pela aplicação de um conjunto de dados denominados de amostras ou conjunto de treinamento. A cada passo do treinamento as amostras são apresentadas à rede. Uma amostra possui um vetor de entrada e uma saída desejada que equivale a um par estímulo-resposta. Apresenta-se aleatoriamente à rede uma amostra e os pesos da rede são ajustados de modo a minimizar o erro entre a saída obtida e a saída desejada. O treinamento é repetido para todo o conjunto de dados até que se obtenha um conjunto de pesos em que o erro global seja o menor possível. Esse erro global considera os erros de todas as amostras, normalmente o critério utilizado é o Erro Quadrático Médio (Mean Squares Error – MSE). Costuma-se dizer que após o término do treinamento a rede construiu um mapeamento de entrada-saída (Haykin, 1994). Após o treinamento é possível testar a capacidade de generalização da rede utilizando amostras que não foram usadas no treinamento da rede (Pasti, 2007).

#### **4.5.2 Aprendizagem não-supervisionada**

Nesse paradigma não existe um supervisor no processo de treinamento para avaliar o desempenho da rede, ou seja, não há medida de erro e as amostras não são rotuladas. Um tipo particularmente importante de aprendizagem não-supervisionada é a aprendizagem auto-organizada (Pasti, 2007).

Esse tipo de rede busca regularidades estatísticas nos dados de entrada. O processo se baseia numa rede com uma camada de neurônios sensoriais que envia as entradas para a próxima camada de neurônios competitivos. Esses neurônios competem entre si para se tornarem ativos e apenas um neurônio será o vencedor para um determinado padrão em um instante  $t$ . Após o processo de treinamento, os neurônios devem ser capazes de identificar regularidades estatísticas nos dados de entrada. Devido a essa característica, esse tipo de aprendizado é muito utilizado em agrupamento de dados (Pasti, 2007).

Basicamente, o processo para se determinar o neurônio vencedor se dá tomando aquele que está mais próximo da amostra de entrada, segundo uma medida de distância ou proximidade. Geralmente utiliza-se a distância Euclidiana. Após a determinação do neurônio vencedor, é feito o ajuste dos pesos.

### **4.5.3 Aprendizagem por reforço**

Nesse paradigma há uma interação sucessiva e direta da rede neural com o ambiente, mas não há um aprendizado supervisionado explícito. A rede tem um objetivo a ser alcançado de modo a se melhorar o desempenho global do aprendizado. Durante a fase de treinamento, a rede interage com o ambiente de acordo com suas ações, recebe uma penalização ou uma recompensa dependendo de seu desempenho. Neste caso, o desempenho é medido através de um escalar que fornece a qualidade da solução encontrada, ou seja, mede o quão bem a rede está se saindo em sua tarefa. O valor deste escalar é baseado na interação da rede com o ambiente através dos valores dos pesos, a resposta da rede e recompensas (de Castro, 2006). Fazendo uma analogia, é como um processo de aprendizado por tentativa e erro no qual com o passar do tempo nos aperfeiçoamos em determinadas tarefas por meio de uma “punição” ao errar e uma “recompensa” ao acertar (Pasti, 2006).

#### **4.5.4 Modos de treinamento**

O treinamento pode ser feito de dois modos, aquele em que as correções nos pesos sinápticos e limiares de ativação só são feitas após uma época, isto é, após a apresentação de todo o conjunto de treinamento. Este é chamado de treinamento por batelada. O outro é o treinamento por padrão (ou incremental) no qual a atualização dos parâmetros da rede neural é feita após cada iteração, ou seja, após a passagem de cada padrão. Na grande maioria das vezes e, nesse trabalho, escolhe-se o tipo de treinamento supervisionado e o modo batelada (Schmitz, 2006).

#### **4.6 Treinamento das RNAs**

No processo de treinamento, os dados de entrada-saída são fornecidos e a rede é ajustada com pesos para minimizar o erro de predição. Nessa fase, seguindo o algoritmo de treinamento escolhido, serão ajustados os pesos das conexões. É importante considerar também, nessa fase, alguns aspectos tais como a inicialização da rede, o tempo de treinamento, o número de neurônios na camada escondida, o número de neurônios em cada camada e o número de dados de treinamento. Essas escolhas não são triviais, na verdade são muito importantes para se conseguir bons resultados nas redes.

Uma boa escolha dos valores iniciais dos pesos da rede pode diminuir o tempo necessário para o treinamento. Normalmente, os valores iniciais dos pesos vêm de um gerador de números randômicos. A escolha errada desses pesos pode levar a uma saturação prematura da rede, levando a RNA a convergir em mínimos locais e que o erro seja grande.

Quanto ao tempo de treinamento, vários fatores podem influenciar a sua duração, porém sempre será necessário utilizar algum critério de parada. O treinamento deve ser interrompido quando a rede apresentar uma boa capacidade de generalização e quando a taxa de erro for suficientemente pequena, ou seja, menor que um erro admissível. Assim, deve-se encontrar um ponto ótimo de parada com erro mínimo e capacidade de generalização máxima. Outro fator que influencia o critério de parada é número máximo de ciclos de iteração (*epochs*).

O número de neurônios nas camadas de entrada e saída da rede normalmente é definido por especificações de problemas externos. Por exemplo, se existem quatro variáveis externas para serem usadas como entradas existem quatro neurônios de entrada, correspondentemente o mesmo acontece com a camada de saída da rede.

Segundo Sen e Yang (2000), quanto à questão do número de neurônios na camada escondida, este parâmetro é mais difícil de se determinar pois não existe uma base teórica consolidada para isso. Mas existem diversas pesquisas nessa área. Hecht-Nielsen, citado em Sen e Yang (2000), mostra que somente uma camada escondida é suficiente para modelar a relação de entrada-saída conquanto que a camada intermediária tenha  $2J_1 + 1$  neurônios, sendo que  $J_1$  é o número de neurônios na camada de entrada. Mas para problemas que possuam um grande número de dados de entrada, os neurônios na camada escondida podem se tornar excessivos. Um grande número de neurônios na camada intermediária além de tornar a rede lenta tanto para o treinamento quanto para a predição, pode causar também o *overfitting* (sobre-ajuste) da rede.

Outra sugestão é dada por Roger e Jenkins, citado em Sen e Yang (2000) é:

$$N_t = 1 + N_n \left( \frac{J_1 + J_I + 1}{J_I} \right) \quad i = 1, 2, \dots, I \quad (4.1)$$

Na equação acima,  $N_t$  é o número de dados de treinamento,  $N_n$  é o número total de neurônios internos da rede e  $J_i$  é o número de neurônios na camada  $i$ . Se  $N_t$ ,  $J_1$  e  $J_I$  são conhecidos no problema, a Equação (4.1) determina o número mínimo de neurônios sugerido nas camadas internas. Também, se  $N_n$ ,  $J_1$  e  $J_I$  são conhecidos, a equação dá o valor de  $N_t$ . O número de dados usados deve ser maior do que o sugerido pela equação para assegurar valores adequados dos pesos e bias no processo de treinamento (Sen e Yang, 2000).

O conjunto de dados de treinamento deve ser suficientemente extenso para garantir a convergência e a estabilidade da rede. Além disso, esses dados devem ser igualmente distribuídos sobre toda a faixa de valores que se deseja predizer (Schmitz, 2002).

Outra forma de se escolher o número ótimo de neurônios na camada escondida é sugerida por Dan Foresee e Hagan (1997) em que o número de neurônios na camada

intermediária é escolhido através do número de parâmetros efetivo (?). O ? é a medida de como muitos parâmetros na rede neural são efetivamente usados na redução da função erro. Este parâmetro varia de 0 a  $N$ . Se ? é muito próximo do número atual de parâmetros ( $N$ ), a RNA não deve ser grande o suficiente para representar a rede. Nesse caso, simplesmente acrescentam-se mais neurônios na camada escondida e treina novamente a rede. Se a nova rede (maior) tem o mesmo valor final de ?, então a outra rede (menor) era grande o suficiente. Neste caso, adicionam-se mais camadas escondidas na rede (Dan Foresee e Hagan, 1997). Deve ser comparado também o valor da soma dos erros quadráticos (SSE), Eq. (4.3), que se torna constante à medida que o número de neurônios da camada escondida aumenta. O método de Dan Foresee e Hagan só é válido para redes que utilizam como método de regularização da rede a regularização Bayesiana.

$$erro = (t_{exp} - t_{calc}) \quad (4.2)$$

$$SSE_{teste} = \sum (erro)^2 \quad (4.3)$$

Na equação acima,  $t_{exp}$  é o valor experimental (desejado) e  $t_{calc}$  é o valor calculado pela rede neural.

## 4.7 Algoritmos de treinamento

Nessa seção serão apresentados os algoritmos de treinamento e regularização usados nessa dissertação.

### 4.7.1 Algoritmo *backpropagation*

O algoritmo *backpropagation* compara a saída desejada com a saída calculada pela rede, calculando o erro. O erro é propagado a partir da camada de saída até a camada de entrada e os pesos das conexões das unidades das camadas internas vão sendo modificados

conforme o erro retro-propagado. O objetivo do algoritmo *backpropagation* é ajustar os parâmetros da rede para minimizar o mean squared error (MSE) (Hagan *et al.*, 1996), que é calculado através da Equação (4.4). Maiores detalhes do algoritmo *backpropagation*, podem ser encontrados em: Haykin (1994), Hagan *et al.* (1996) e de Castro (2006).

$$MSE = \frac{SSE_{teste}}{n} \quad (4.4)$$

Do qual  $n$  é o número de dados usados, neste caso, no conjunto de teste da rede.

#### 4.7.2 Algoritmo de Levenberg-Marquardt

O algoritmo de Levenberg-Marquardt é uma variação do método de Newton, que é usado para minimizar funções não lineares. Esse algoritmo é usado junto com o algoritmo *backpropagation*, pois acelera a convergência da rede (Hagan *et al.*, 1996). Esse algoritmo é visto com maiores detalhes em: Haykin (1994), Hagan *et al.* (1996) e de Castro (2006).

#### 4.7.3 Regularização Bayesiana

O objetivo do treinamento de uma rede neural é fazer não só que essa apresente pequenos erros para o conjunto de treinamento, mas também que responda bem para entradas ainda desconhecidas. Quando uma rede neural é capaz de responder adequadamente a uma nova entrada, diz-se que ela generaliza bem. Porém, quando a rede representa bem apenas o conjunto usado no treinamento, ela não é capaz de generalizar e diz-se que esta memorizou os padrões de treinamento começando o *overfitting* (sobre-ajuste), que é um dos problemas que ocorrem durante o treinamento da rede. A memorização é característica de redes superdimensionadas. O emprego de RNAs com um número de neurônios grande o suficiente para ajustar os dados adequadamente, mas que não tenha força suficiente para memorizar é o ideal. Mas a obtenção desse número exato de

neurônios é difícil. Por esse motivo surgiram outros métodos para melhorar a generalização das redes neurais, chamados métodos de regularização (Dan Foresee e Hagan, 1997).

A regularização Bayesiana minimiza a combinação linear dos erros quadráticos e pesos. Ela modifica a combinação linear até que no final do treinamento o resultado da rede esteja com uma boa qualidade de generalização. Essa regularização acontece dentro do algoritmo de Levenberg-Marquardt.

Quando a regularização Bayesiana é usada, os melhores resultados são obtidos se os dados de treinamento são primeiro normalizados na faixa de  $[-1,1]$  (ou regiões similares). Coloca-se nessa faixa ambas entradas e saídas da rede (Dan Foresee e Hagan, 1997).

## 5 Metodologia

Nesse trabalho foi utilizado o pacote de redes neurais do *software* Matlab 7.0. Para o treinamento da rede usou-se a função TRAINBR. TRAINBR é uma função de treinamento que atualiza os pesos e bias usando o algoritmo *backpropagation* baseado no algoritmo de Levenberg-Marquardt e na regularização Bayesiana. As RNAs desenvolvidas nesse trabalho são redes *feedforward* com uma camada escondida e inicialização randômica. Como função de ativação foi usada a função “tansig” para a camada escondida e a função “purelin” para a camada de saída da rede.

As RNAs simuladas nessa dissertação foram avaliadas através dos gráficos gerados pelo Matlab (variável predita versus variável desejada), do R de treinamento e teste, do SSE de treinamento e teste, e do MSE de teste da rede.

Foram implementadas cinco propostas de rede com as seguintes características:

**Rede R1:** Nessa rede foram usados somente os componentes álcoois-água com cinco grupos funcionais UNIFAC e foi implementada a simplificação da proposta 1, ou seja, o cálculo do  $\gamma^8$  através das RNAs.

**Rede R2:** Nessa rede foi implementada a proposta 1, ou seja, RNAs substituindo o modelo UNIFAC no cálculo do  $\gamma$  e foram usados os mesmos componentes e grupos funcionais UNIFAC da rede R1.

**Rede R3:** A rede R3 usa a mesma proposta da rede R2, mas trabalha com diversos componentes (álcoois-água, aldeídos-água, cetonas-água, álcoois-álcoois, álcoois-aldeídos, álcoois-cetonas, aldeídos-cetonas e cetonas-cetonas) com oito grupos funcionais UNIFAC.

**Rede R4:** A rede R4 utiliza a proposta 2, ou seja, o cálculo direto do  $y$  através das RNAs, essa rede trabalha somente com os componentes álcoois-água com cinco grupos funcionais UNIFAC.

**Rede R5:** A rede R5 utiliza a mesma proposta da rede R4, mas trabalha com os mesmos componentes e grupos funcionais UNIFAC da rede R3.

## 6 Resultados e Discussões

Nesse capítulo serão apresentados e analisados os resultados obtidos pelas redes neurais artificiais propostas nessa dissertação.

### 6.1 Projeto da rede R1

A rede R1 foi criada como uma simplificação da proposta 1, já que não usa  $\omega$  e  $\gamma^s$ . Essa rede foi treinada e testada com dados isotérmicos e isobáricos de misturas binárias álcoolis-água. Os grupos UNIFAC escolhidos para entrada na rede foram: CH<sub>3</sub>, CH<sub>2</sub>, CH, OH e H<sub>2</sub>O. Todos os dados ( $x_1$ ,  $y_1$ ,  $\gamma^s$ ,  $P$  e  $T$ ) foram retirados da Coleção DECHEMA.

A rede R1 possui como entradas os grupos UNIFAC do componente 1, grupos UNIFAC do componente 2 e  $T$ , e como saída  $\gamma^s_1$ , vide Figura 6.1. Na Tabela 6.1 são apresentados os componentes e grupos UNIFAC utilizados na rede R1.

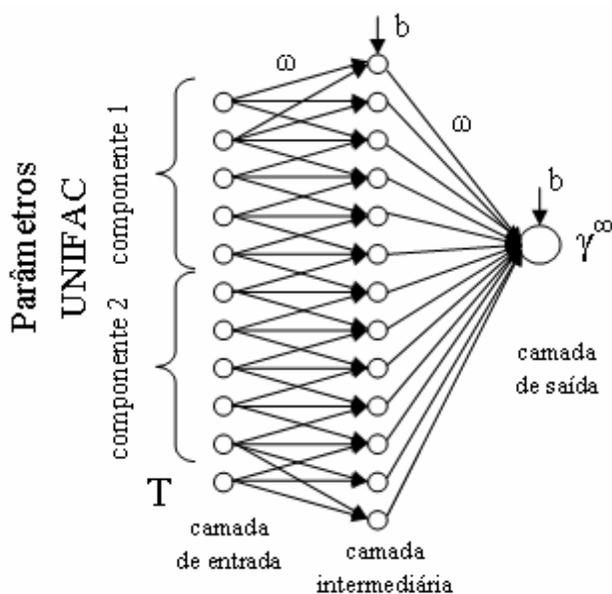


Figura 6.1 - Representação da rede R1.

Tabela 6.1 - Componentes (álcoois-água) e grupos UNIFAC.

Compostos	CH3	CH2	CH	OH	H2O
Água	0	0	0	0	1
Metanol	1	0	0	1	0
Etanol	1	1	0	1	0
1-Propanol	1	2	0	1	0
2-Propanol	2	0	1	1	0
2-Metil-1-Propanol	2	1	1	1	0
1,2-Propanodiol	1	1	1	2	0
1-Butanol	1	3	0	1	0
2-Butanol	2	1	1	1	0
2,3-Butanodiol	2	0	2	2	0
1-Hexanol	1	5	0	1	0
2-Hexanol	1	5	0	1	0
Ciclohexanol	0	5	1	1	0

Testaram-se redes neurais para, inicialmente, todos os componentes escolhidos, onde existiam coeficientes de atividade à diluição infinita com valores de até 71,31; com 130 pontos de treinamento e 36 pontos de teste. O conjunto de teste escolhido foram os dados da mistura binária etanol-água. O treinamento da rede foi considerado bom, mas o conjunto de teste não apresentou o mesmo comportamento. Por conseguinte, restringiram-se os valores do  $\gamma^s$  para até 10, com 100 pontos de treinamento e 20 pontos de teste, com temperaturas de 15 a 239 °C e pressões de 25 a 5001 mmHg. Mesmo assim, o conjunto de teste apresentou valores muito distantes dos retirados da literatura. Uma possível fonte desse problema, é que os coeficientes de atividade à diluição infinita fornecem somente dois pontos da curva de ELV, os pontos extremos da curva, quando  $x = 0$  e  $x = 1$ , o que representa pouca informação para a rede.

A rede R1 foi testada com dois tipos de função de ativação “tansig” e “logsig” na camada escondida e com a função “purelin” na camada de saída. A rede também foi testada com duas camadas de neurônios escondidos, usando as funções tansig-tansig-purelin respectivamente, mas essas alterações não apresentam melhora no desempenho da rede.

### 6.1.1 Resultados da rede R1

Para análise das redes foram usados o somatório dos quadrados dos erros (SSE) do treinamento e teste, o coeficiente de correlação (R) e a média dos erros quadráticos (MSE) dos testes. Os valores do SSE do conjunto de treinamento e o R tanto para o treinamento quanto para o teste são fornecidos pelo Matlab. Já os valores do SSE do conjunto de teste e MSE foram calculados através das Equações (4.2), (4.3) e (4.4).

Na Equação (4.2),  $t_{exp}$  são os dados experimentais retirados da Coleção DECHEMA ou dados calculados a partir das variáveis fornecidas pela Coleção DECHEMA,  $t_{calc}$  são os dados calculados pela rede neural e  $n$  é o número de dados usados para teste na rede.

As Figuras 6.2 e 6.3 mostram os gráficos de treinamento e teste para a rede neural R1 com 20 neurônios na camada escondida. Esse número de neurônios foi escolhido aleatoriamente, pois a rede R1 foi treinada para diversos números de neurônios na camada escondida (5, 10, 15, 20 e 30), mas como o resultado da rede (conjunto de testes) não foi considerado bom, o número de neurônios na camada escondida não foi determinado. A Tabela 6.2 apresenta os dados de treinamento e teste obtidos pela rede (SSE, R e MSE).

A linha de ajuste que consta na legenda das Figuras 6.2 e 6.3 é a regressão linear do conjunto de dados calculado pelo Matlab. Em um caso ideal, considerando a equação  $y = ax + b$  e R, onde  $a$  é o coeficiente angular,  $b$  é o coeficiente linear,  $x$  é o eixo das abscissas,  $y$  é o eixo das coordenadas e R é o coeficiente de correlação;  $a$  é igual a 1,  $b$  é igual a 0 e R é igual a 1. Os dados experimentais são os dados retirados da Coleção DECHEMA. A linha predito = real é a linha diagonal desenhada pelo Matlab que é usada para comparação com a linha de ajuste dos dados.

Tabela 6.2 - Dados de treinamento e teste obtidos pela rede R1 com 20 neurônios na camada escondida.

Parâmetros da rede	?
SSE treinamento	0,300
R treinamento	0,987
SSE teste	15,780
R teste	0,859
MSE teste	0,789

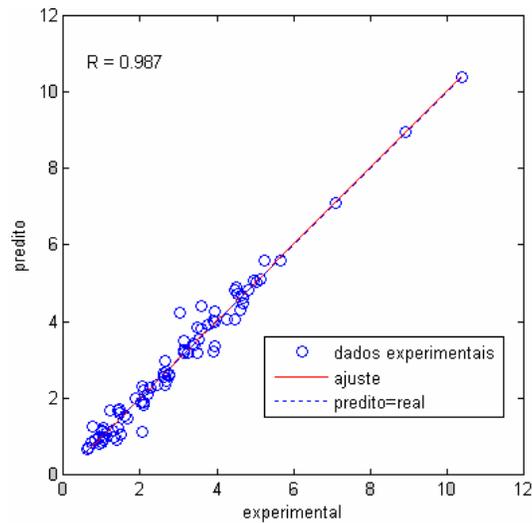


Figura 6.2 - Comparação entre valores experimentais do ? com valores calculados pela rede neural R1 no treinamento, 20 neurônios na camada escondida.

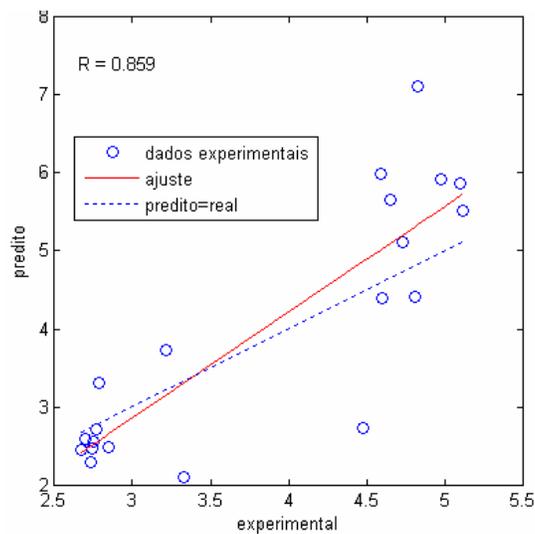


Figura 6.3 - Comparação entre valores experimentais do ? com valores calculados pela rede neural R1 no teste, 20 neurônios na camada escondida.

### 6.1.2 Conclusões parciais da rede R1

Através das Figuras 6.2 e 6.3 e Tabela 6.2 percebe-se que a rede R1 realmente não alcançou bons resultados nos testes. A rede obteve um bom ajuste para os dados de treinamento, o SSE de treinamento foi baixo, o R foi próximo de 1,0 e o gráfico do treinamento apresentou seus pontos ao longo da diagonal.

Mas para os dados de teste ocorreu o contrário, o gráfico obteve um ajuste muito ruim, pois apresentou pontos dispersos pelo gráfico distantes da diagonal, SSE alto comparado com o do treinamento, R distante de 1,0 e MSE alto. O MSE foi considerado alto, pois 0,789 é um número considerável quando comparado com coeficientes de atividade à diluição infinita que estão na faixa de 0,661 a 10,420.

O estudo da rede R1 foi então abandonado e voltou-se novamente para a proposta 1 dessa dissertação que usa  $\gamma$  para prever o ELV de misturas binárias não ideais.

## 6.2 Projeto da rede R2

A rede R2 é a proposta 1 sugerida nessa dissertação, ou seja, a substituição do método UNIFAC no cálculo do  $\gamma$  por RNAs.

As entradas da rede são as mesmas da rede R1, com a adição de  $x_1$ , pois  $\gamma$  depende da composição da mistura, Figura 3.5, e como saída da rede  $\gamma_1$ , vide Figura 6.4. Foram usados os mesmos componentes e grupos da Tabela 6.1 e como conjunto de teste foram usados os dados da mistura binária etanol-água. O grupo etanol-água foi escolhido por apresentar dados em toda a faixa de temperatura e pressão estudados e por ser um sistema de relevante importância. Os binários foram trocados no arquivo de dados da rede, exemplificado na Tabela 6.3, para que a rede não vicie, já que a água ficaria sempre na posição do componente 2. A rede R2 tem 882 pontos de treinamento e 200 pontos de teste, com temperaturas de 15 a 215 °C e pressões de 1,29 a 5001 mmHg.

Tabela 6.3 - Exemplo de troca de binários.

Comp. 1	Comp. 2	Grupos UNIFAC comp. 1					Grupos UNIFAC comp. 2					T (K)	P (mmHg)	x1	?1
		CH3	CH2	CH	OH	H2O	CH3	CH2	CH	OH	H2O				
etanol	água	1	0	0	1	0	0	0	0	1	298	41,59	0,1	1,611	
água	etanol	0	0	0	0	1	1	0	0	1	298	117,07	0,1	1,457	

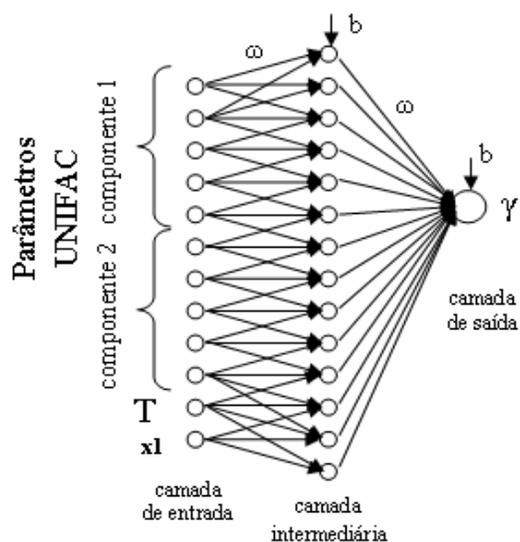


Figura 6.4 – Representação da rede R2.

Uma vez que a rede R2 trabalha com ? que é dependente da composição, os dados de entrada tiveram que ser expandidos. Para o cálculo do ? nas diversas composições ( $x_1 = 0,1; 0,2; \dots; 0,9$ ) foram criados dois programas, um para o cálculo da temperatura de bolha (*Bolha T*), e outro para o cálculo da pressão de bolha (*Bolha P*). No cálculo do *Bolha T* e *Bolha P* foi usada a equação de Antoine indicada pelo DECHEMA e seus parâmetros foram retirados do mesmo.

$$\log P^{sat} = A - \frac{B}{T + C} \quad (6.1)$$

Sendo que  $A$ ,  $B$  e  $C$  são os parâmetros de Antoine,  $P^{sat}$  em mmHg e  $T$  em °C.

### 6.2.1 Temperatura de bolha (*Bolha T*)

O algoritmo da temperatura de bolha é feito para dados isobáricos ( $P$  constante). Para o cálculo da temperatura de bolha foram feitas as seguintes etapas:

- Estima-se uma temperatura ( $T$ );
- Calcula-se  $P_1^{sat}$  e  $P_2^{sat}$  através da equação de Antoine (Eq. 6.1);
- Calcula-se o  $\gamma_1$  e  $\gamma_2$  para cada  $x_1$  ( $x_1 = 0,1; 0,2; \dots; 0,9$ );
- Testa-se o valor do  $T$  inicial através da equação de Raoult estendida para a mistura binária (Eq. 6.5). Se  $b = 0$  o valor do  $T$  estimado foi correto, ou seja,  $T$  estimado =  $T$  de bolha, senão volta-se ao início do algoritmo estimando-se um novo  $T$ .

A Equação (6.5), que é usada para o teste dos algoritmos *Bolha T* e *Bolha P* foi deduzida através da soma das Equações (3.2) do componente 1 e 2, resultando na Equação (6.2). Essa, por sua vez, foi simplificada através da Equação (6.3), resultando nas Equações (6.4) e (6.5).

$$\begin{aligned} & y_1 P = x_1 g_1 P_1^{sat} \\ + & y_2 P = x_2 g_2 P_2^{sat} \\ \hline & (y_1 + y_2) P = x_1 g_1 P_1^{sat} + x_2 g_2 P_2^{sat} \end{aligned} \quad (6.2)$$

Como para misturas binárias:

$$y_1 + y_2 = 1 \quad (6.3)$$

A Eq. (6.2) se torna:

$$x_1 g_1 P_1^{sat} + x_2 g_2 P_2^{sat} - P = 0 \quad \text{condição do bolha } T \quad (6.4)$$

$$x_1 g_1 P_1^{sat} + x_2 g_2 P_2^{sat} - P = b \quad \text{teste do bolha } T \quad (6.5)$$

O cálculo do  $\gamma$  foi feito utilizando a equação indicada pelo DECHEMA como a que obteve o melhor ajuste, foram utilizadas as equações de Margules, van Laar, Wilson e NRTL. Somente a equação UNIQUAC não foi utilizada devido a sua complexidade, quando o DECHEMA indicava essa equação como melhor ajuste, escolhia-se uma dessas outras equações indicadas acima. A Figura 6.5 mostra o algoritmo da temperatura de bolha (*Bolha T*).

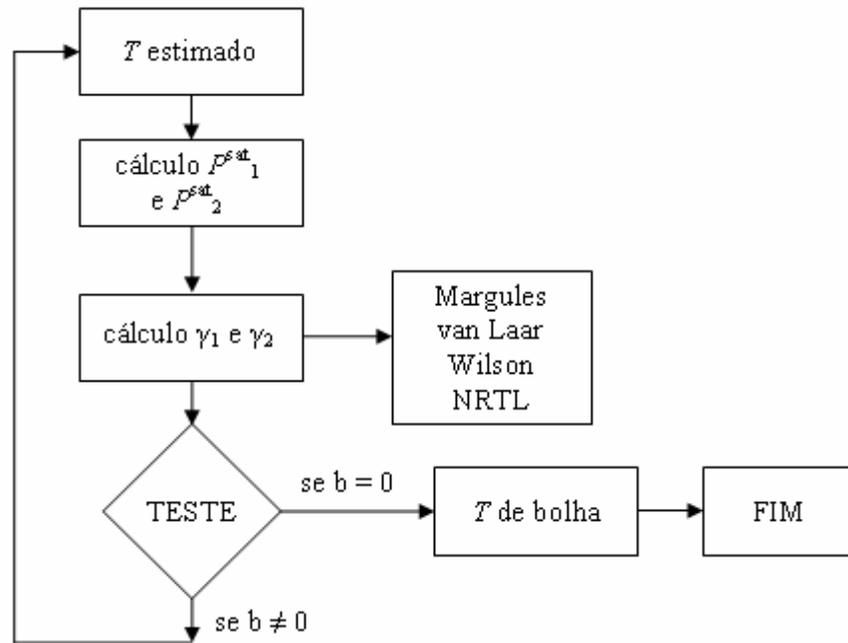


Figura 6.5 – Algoritmo da temperatura de bolha (*Bolha T*).

### 6.2.2 Pressão de bolha (*Bolha P*)

O cálculo da pressão de bolha é feito para dados isotérmicos ( $T$  constante). Para o cálculo da pressão de bolha foram feitas as seguintes etapas:

- Calcula-se o  $P_1^{sat}$  e  $P_2^{sat}$  pela equação de Antoine;
- Calcula-se  $\gamma_1$  e  $\gamma_2$ ;
- Estima-se uma pressão ( $P$ );

- Testa-se o valor do  $P$  inicial através da equação de Raoult estendida para a mistura binária (Eq. 6.5). Se  $b = 0$  o valor do  $P$  estimado foi correto, ou seja,  $P$  estimado =  $P$  de bolha, senão volta-se ao início do algoritmo estimando-se um novo  $P$ .

A Figura 6.6 mostra o algoritmo da pressão de bolha (*Bolha P*).

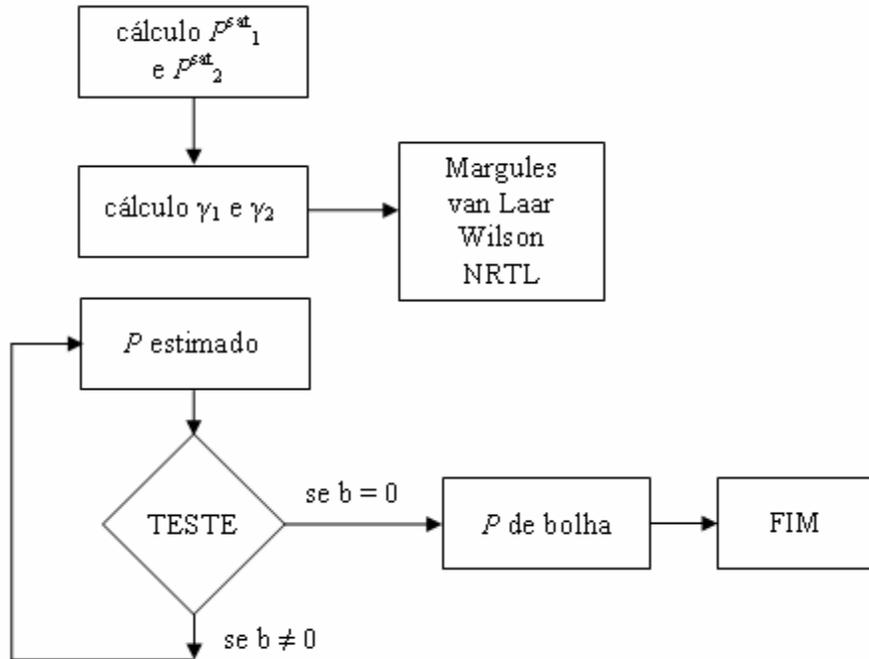


Figura 6.6 – Algoritmo da pressão de bolha (*Bolha P*).

### 6.2.3 Escolha do número de neurônios na camada escondida

Segundo Sen e Yang (2000), não existe uma teoria consolidada para a escolha do número de neurônios na camada escondida, mas existem diversos trabalhos na área. Nessa dissertação foram utilizados alguns métodos para análise desse parâmetro da rede. Na seção 3.2.4 foram mencionados alguns deles, a Tabela 6.4 mostra os métodos estudados, as equações utilizadas e o número de neurônios na camada escondida calculados pelos mesmos.

A rede R2 foi simulada de 5 em 5 neurônios (5, 10, 15, ..., 40) até o valor de 40 neurônios na camada escondida. Na faixa onde os valores começaram a ficar constantes, a rede foi simulada variando o número de neurônios de uma em uma unidade (25, 26, 27, 28, 29 e 30). Cada neurônio foi simulado pelo menos três vezes, ao final dessa série de treinamento foi feita a média do SSE e do  $\rho$ , estas médias foram os valores considerados nas análises e gráficos. Nas Figuras 6.7 e 6.8, o número de neurônios escolhido, no caso, 28 neurônios na camada intermediária, está representado em destaque por um triângulo.

O método de Dan Foresee e Hagan (1997) foi o escolhido para sugerir o número de neurônios na camada escondida, os outros métodos são mostrados apenas a título de comparação.

Tabela 6.4 - Métodos utilizados para encontrar o número de neurônios na camada escondida para a rede R2.

<b>Correlações</b>	<b>Equações</b>	<b>Número de Neurônios na Camada Escondida</b>
Hecht-Nielsen	$N_n = 2J_1 + 1$	<b>25</b>
Rogers e Jenkins	$N_t = 1 + N_n \left( \frac{J_1 + J_t + 1}{J_t} \right)$	<b>63</b>
Dan Foresee e Hagan	? ? $N$ , ? e SSE do treinamento constantes	<b>28</b>

Na Tabela 6.4,  $J_1$  é o número de neurônios na camada de entrada,  $N_n$  é o número de neurônios na camada escondida,  $N_t$  é o número de dados de treinamento e  $J_t$  é o número de neurônios na camada de saída. Para a rede R2,  $J_1$  é igual a 12,  $J_t$  é igual a 1 e  $N_t$  é igual a 882. Quando o número de neurônios na camada escondida não é inteiro, arredonda-se esse número para cima.

As Figuras 6.7 e 6.8 mostram o  $\rho$  e o SSE, respectivamente, versus o número de neurônios na camada escondida.

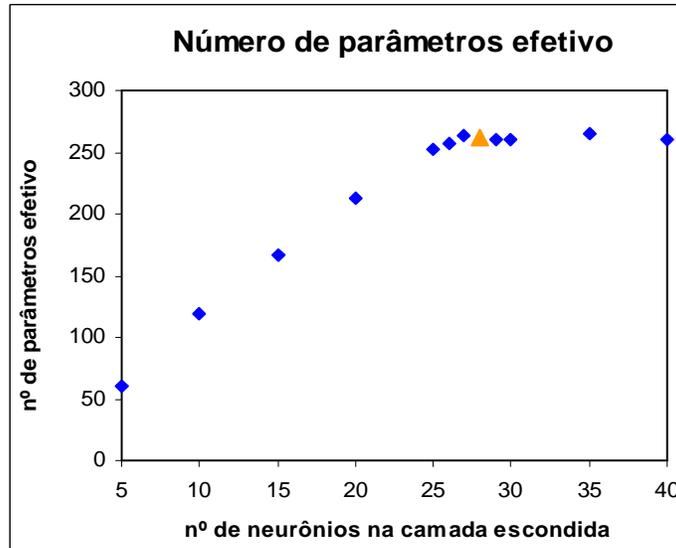


Figura 6.7 - Gráfico do número de parâmetros efetivos pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 28.

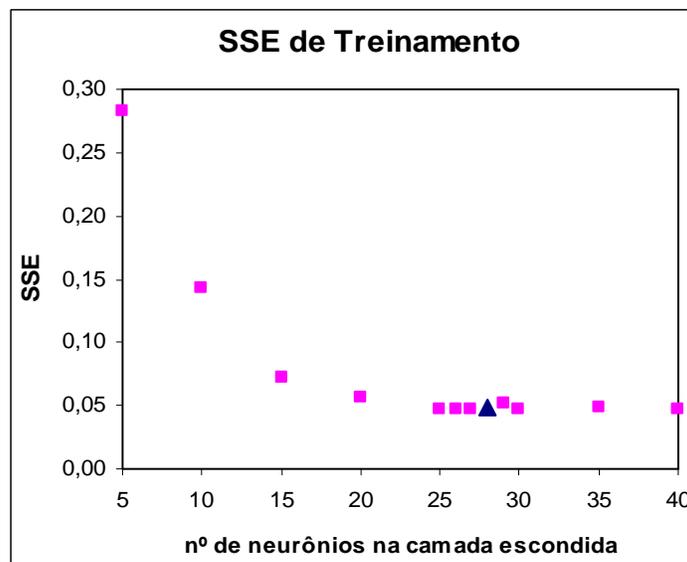


Figura 6.8 - Gráfico do SSE pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 28.

Analisando-se a Tabela 6.4 nota-se que o primeiro e o último método apresentaram resultados similares, enquanto o segundo método resultou em um valor bem diferente e também muito mais alto do que os demais. Um número de neurônios de 63 é considerado, para esse trabalho, um valor alto, pois torna o treinamento da rede muito lento e também aumenta muito o número de parâmetros (pesos e bias) utilizados pela RNA. Como mencionado anteriormente, o primeiro e o segundo método foram usados somente a título

de comparação, sendo considerado o valor do terceiro método como o número de neurônios na camada escondida escolhido, portanto, para a rede R2, 28 neurônios na camada intermediária. O número de 28 neurônios foi escolhido por estar na faixa onde o SSE e o  $\rho$  se tornam praticamente constantes e o  $N$  é diferente de  $\rho$  ( $N = 393,0$  e  $\rho = 262,1$ ).

#### 6.2.4 Resultados da rede R2

A Tabela 6.5 apresenta os valores obtidos pela rede R2 ao ser simulada com 28 neurônios na camada escondida. As Figuras 6.9 e 6.10 mostram os gráficos de treinamento e teste da rede.

Tabela 6.5 - Dados de treinamento e teste obtidos pela rede R2 com 28 neurônios na camada escondida.

Parâmetros da rede	$\rho$
SSE treinamento	0,061
R treinamento	0,999
SSE teste	4,994
R teste	0,961
MSE teste	0,025

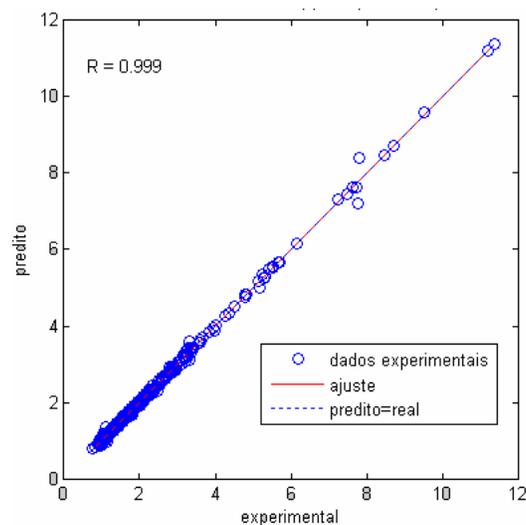


Figura 6.9 - Comparação entre valores experimentais do  $\rho$  com valores calculados pela rede neural R2 no treinamento, 28 neurônios na camada escondida.

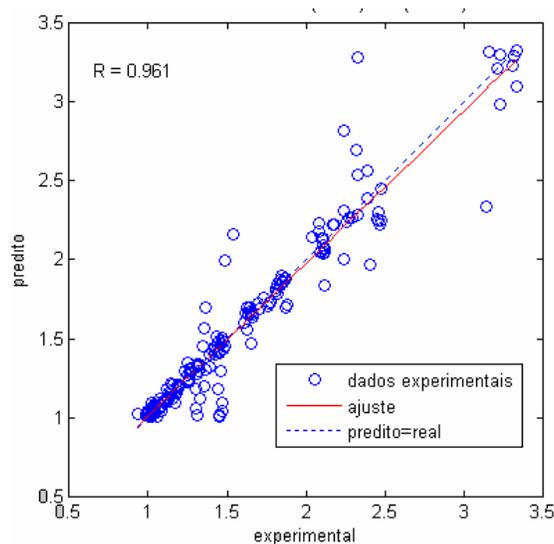


Figura 6.10 - Comparação entre valores experimentais do  $\eta$  com valores calculados pela rede neural R2 no teste, 28 neurônios na camada escondida.

### 6.2.5 Conclusões parciais da rede R2

As Figuras 6.9 e 6.10 mostram que a rede R2 teve um bom comportamento, tanto para o treinamento quanto para o teste. No gráfico do teste alguns valores foram muito distantes da diagonal, o que denota um grande erro, mas foram valores pontuais não podendo-se desconsiderar o bom desempenho da rede.

Os valores da Tabela 6.5 mostram que o treinamento da rede foi muito bom, pois esta apresentou SSE baixo e R próximo de 1,0. Já no teste o SSE não foi tão baixo assim, mas o MSE indicou um valor pequeno (0,025) pois este valor não tem influência considerável nos coeficientes de atividade, que variam de 0,935 a 3,342, lembrando que o MSE é a média dos erros quadráticos.

É de se considerar que os valores com maiores erros não são de componentes específicos e que estes ocorrem na maioria das vezes perto dos extremos, quando  $x_1$  igual a 0,1 e  $x_1$  igual a 0,9. Isso acontece porque geralmente essa é a região de maior não idealidade do sistema.

### 6.3 Projeto da rede R3

A rede R3 representa também a proposta 1 sugerida nessa dissertação, Figura 6.4. Visto que a rede anterior, rede R2, foi capaz de prever o ELV, expandiu-se a rede R3 para novos componentes e grupos funcionais UNIFAC. Além dos binários álcoois-água, têm-se também as misturas: aldeídos-água, cetonas-água, álcoois-álcoois, álcoois-aldeídos, álcoois-cetonas, aldeídos-cetonas e cetonas-cetonas. Os grupos UNIFAC são, além dos já utilizados nas redes anteriores CH<sub>3</sub>, CH<sub>2</sub>, CH, OH e H<sub>2</sub>O, foram adicionados os grupos CHO, CH<sub>3</sub>CO e CH<sub>2</sub>CO. A Tabela 6.6 apresenta os grupos UNIFAC e componentes usados na rede R3.

Tabela 6.6 - Componentes e grupos UNIFAC usados na rede R3.

Compostos	CH3	CH2	CH	OH	H2O	CHO	CH3CO	CH2CO
Água	0	0	0	0	1	0	0	0
Metanol	1	0	0	1	0	0	0	0
Etanol	1	1	0	1	0	0	0	0
1-Propanol	1	2	0	1	0	0	0	0
2-Propanol	2	0	1	1	0	0	0	0
2-Metil-1-Propanol	2	1	1	1	0	0	0	0
1,2-Propanodiol	1	1	1	2	0	0	0	0
1-Butanol	1	3	0	1	0	0	0	0
2-Butanol	2	1	1	1	0	0	0	0
2,3-Butanodiol	2	0	2	2	0	0	0	0
1-Hexanol	1	5	0	1	0	0	0	0
2-Hexanol	1	5	0	1	0	0	0	0
Ciclohexanol	0	5	1	1	0	0	0	0
Acetaldeído	1	0	0	0	0	1	0	0
Acetona	1	0	0	0	0	0	1	0
2-Butanona	1	1	0	0	0	0	1	0
Ciclohexanona	0	4	0	0	0	0	0	1
Ciclopropil Metil Cetona	0	2	1	0	0	0	1	0
Isopropil Metil Cetona	2	0	1	0	0	0	1	0
3-Pentanona	2	1	0	0	0	0	0	1
4-Metil-2-Pentanona	2	1	1	0	0	0	1	0

Butiraldeído	1	2	0	0	0	1	0	0
Isobutiraldeído	2	0	1	0	0	1	0	0

As entradas e saídas da rede são as mesmas da rede R2, mas agora tem-se oito grupos UNIFAC para cada componente, ao invés de apenas cinco usados nas redes anteriores.

Inicialmente foram feitos testes com componentes aleatórios no conjunto de teste, mas o desempenho obtido pela rede foi considerado ruim. Por conseguinte, foram feitos novos testes, com componentes que representassem todos os grupos estudados (álcoois-água, aldeídos-água, cetonas-água, álcoois-álcoois, álcoois-aldeídos, álcoois-cetonas, aldeídos-cetonas, cetonas-cetonas) no conjunto de teste. Foram escolhidos os seguintes componentes como grupo de teste da rede R3: acetaldeído-água, metanol-etanol, acetona-metanol, acetona-2-butanona, 1-butanol-água, 2-butanol-água. A rede R3 possui 3240 dados de treinamento e 522 dados de teste, com temperaturas de 10 a 215 °C e pressões de 0,74 a 7109 mmHg.

### 6.3.1 Escolha do número de neurônios na camada escondida

A Tabela 6.7 mostra o número de neurônios na camada escondida encontrados por três correlações para a rede R3.

Tabela 6.7 - Métodos utilizados para encontrar o número de neurônios na camada escondida para a rede R3.

<b>Correlações</b>	<b>Nº de Neurônios na Camada Escondida</b>
Hecht-Nielsen	<b>37</b>
Rogers e Jenkins	<b>162</b>
Dan Foresee e Hagan	<b>57</b>

Na tabela acima,  $J_1$  é igual a 18,  $J_1$  é igual a 1 e  $N_t$  é igual a 3240.

A rede R3 foi simulada de 5 em 5 neurônios até o valor de 70 neurônios na camada escondida. Na faixa onde os valores começaram a se estabilizar a rede foi simulada variando o número de neurônios de uma em uma unidade (55, 56, 57 e 58). Cada neurônio foi simulado pelo menos três vezes, ao final dessa série de treinamento foi feita a média do SSE e do  $\rho$ , estas médias foram os valores considerados nas análises e gráficos. Nas Figuras 6.11 e 6.12, o número de neurônios escolhido, no caso, 57 neurônios na camada intermediária, está representado em destaque por um triângulo.

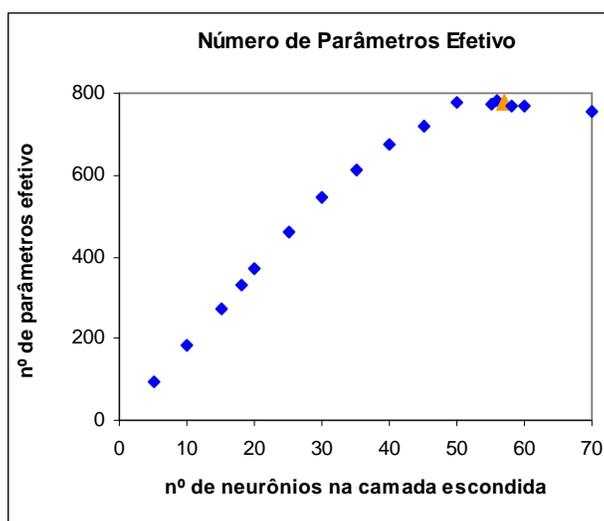


Figura 6.11 - Gráfico do  $\rho$  pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 57.

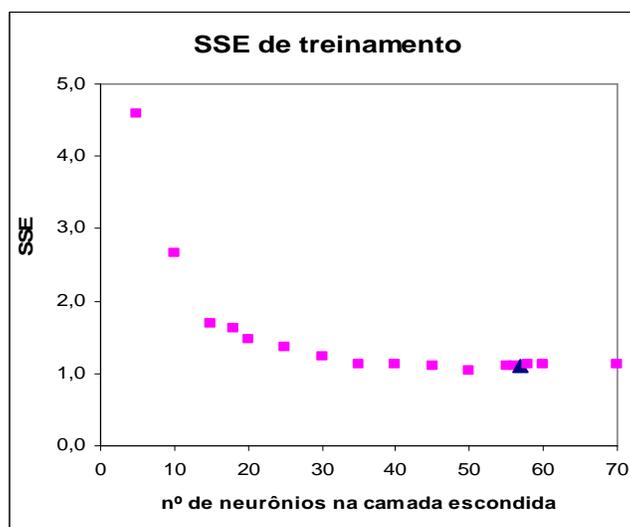


Figura 6.12 - Gráfico do SSE pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 57.

Analisando-se a Tabela 6.7 nota-se que os três resultados dos números de neurônios na camada escondida são muito divergentes e que todos possuem valores muito altos. O método de Rogers e Jenkins obteve um número extremamente alto, tornando inviável a sua aplicação. O valor do número de neurônios do método de Dan Foresee e Hagan também é considerado um valor alto, mas como esse foi considerado o método base dessa dissertação, seu valor foi o utilizado nas simulações da rede R3. O número de 57 neurônios foi o escolhido como ótimo por estar na região onde o SSE e o  $\rho$  são praticamente constantes e o  $N$  é diferente de  $\rho$  ( $N = 1141,0$  e  $\rho = 777,1$ ).

### 6.3.2 Resultados da rede R3

Na Tabela 6.8 são apresentados os valores obtidos pela rede R3 ao ser simulada com 57 neurônios na camada escondida. As Figuras 6.13 e 6.14 representam os gráficos de treinamento e teste da rede.

Tabela 6.8 - Dados de treinamento e teste obtidos pela rede R3 com 57 neurônios na camada escondida.

<b>Parâmetros da rede</b>	<b><math>\rho</math></b>
SSE treinamento	1,142
R treinamento	0,992
SSE teste	107,9
R teste	0,907
MSE teste	0,207

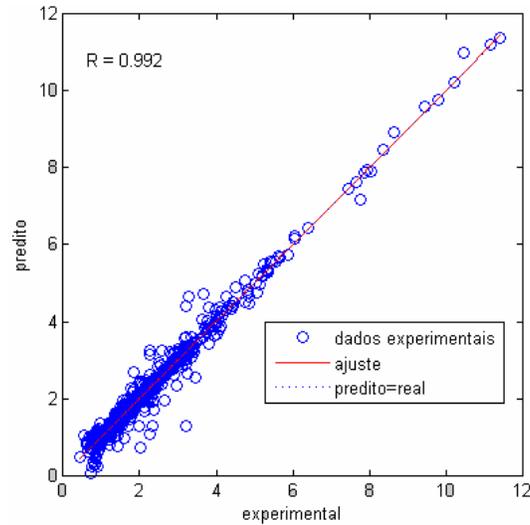


Figura 6.13 - Comparação entre valores experimentais do ? com valores calculados pela rede neural R3 no treinamento, 57 neurônios na camada escondida.

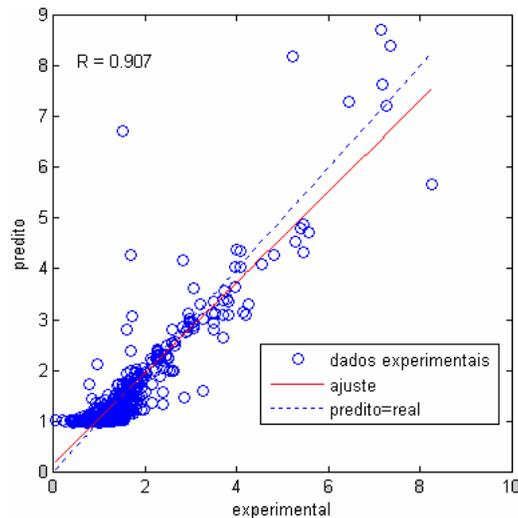


Figura 6.14 - Comparação entre valores experimentais do ? com valores calculados pela rede neural R3 no teste, 57 neurônios na camada escondida.

A Figura 6.13, referente ao treinamento da rede R3 teve um bom desempenho com valor de R próximo de 1,0 e pontos ao longo da reta de ajuste. Os pontos descritos como dados experimentais nas Figuras 6.13 e 6.14 não são experimentais no caso da rede R3, e sim dados calculados a partir de valores retirados da Coleção DECHEMA, algoritmos *Bolha T* e *Bolha P*. O SSE do treinamento pode ser considerado alto se for comparado com o valor do SSE de treinamento da rede R2 (0,061), mas vale lembrar que a rede R2 é uma RNA muito menor do que a rede R3, tanto em número como em tipos de componentes,

número de grupos UNIFAC e número de dados de treinamento e teste, sendo portanto, esperado o aumento do erro, já que a rede R3 é muito mais complexa do que as redes estudadas anteriormente.

Os resultados da Figura 6.14, referente aos dados de teste, não foram considerados bons, pois a reta de ajuste está distante da diagonal, além de se ter vários pontos dispersos no gráfico. O SSE de teste e o MSE foram considerados altos para valores do coeficiente de atividade que variam de 0,203 a 8,265, (MSE = 0,207).

Notou-se ao simular a rede R3 para diferentes números de neurônios, que a mesma obtia um desempenho melhor para números menores de neurônios na camada escondida. Simulou-se então a rede R3 para 18 neurônios na camada intermediária, esse número foi escolhido por ser o mesmo número de entradas da rede ( $N = 361,0$  e  $? = 332,3$ ).

Tabela 6.9 - Dados de treinamento e teste obtidos pela rede R3 com 18 neurônios na camada escondida.

Parâmetros da rede	?
SSE treinamento	2,260
R treinamento	0,984
SSE teste	78,81
R teste	0,929
MSE teste	0,151

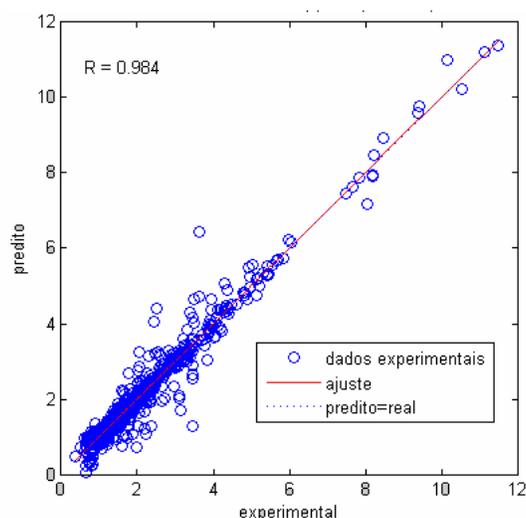


Figura 6.15 - Comparação entre valores experimentais do ? com valores calculados pela rede neural R3 no treinamento, 18 neurônios na camada escondida.

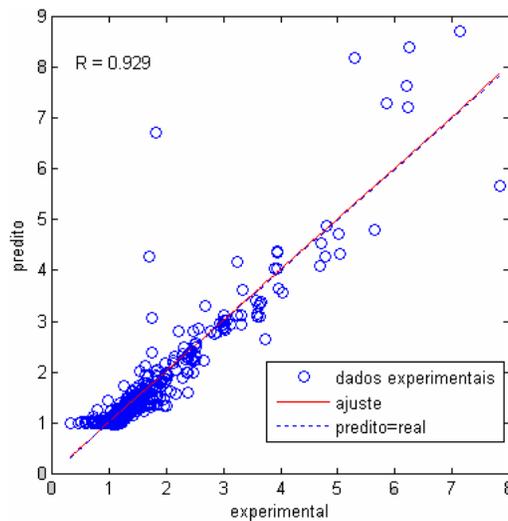


Figura 6.16 - Comparação entre valores experimentais do ? com valores calculados pela rede neural R3 no teste, 18 neurônios na camada escondida.

As Figuras 6.15 e 6.16 para 18 neurônios mostraram um desempenho melhor da rede R3 do que as Figuras 6.13 e 6.14 para 57 neurônios. Os resultados da Figura 6.15, referente ao treinamento da rede para 18 neurônios, apresentaram um desempenho um pouco pior do que os do gráfico de treinamento para 57 neurônios, o SSE do treinamento foi maior, mas mesmo assim o desempenho da rede no treinamento foi considerado bom.

Os resultados da Figura 6.16, referente ao teste da rede para 18 neurônios, foram considerados melhores do que os do gráfico para 57 neurônios. Embora tenha vários pontos dispersos pelo gráfico, a reta de ajuste ficou ao longo da diagonal e o SSE de teste diminuiu o seu valor, mesmo assim, este ainda é considerável. O mesmo ocorreu para o MSE, que ainda é considerado alto ( $MSE = 0,151$ ), visto que os valores dos coeficientes de atividade estão na faixa de 0,311 a 7,858.

### 6.3.3 Conclusões parciais da rede R3

Embora o SSE do teste e o MSE sejam considerados grandes e a Figura 6.16 tenha pontos dispersos da reta de ajuste, o desempenho da rede R3 foi considerado razoável. Já que é uma rede que possui muitos dados de treinamento, componentes com

comportamentos muito não lineares e com faixas de coeficiente de atividade muito diferentes, podendo ser considerada uma rede muito complexa.

É notável também que a simulação com 18 neurônios na camada escondida apresentou um melhor desempenho da rede do que a simulação com 57 neurônios.

## 6.4 Projeto da rede R4

A rede R4 representa a proposta 2 dessa dissertação, ou seja, o cálculo direto do  $y$ . Para isso, primeiramente acrescentou-se a pressão ( $P$ ) do sistema e a composição do componente 1 na fase líquida ( $x_1$ ) na entrada da rede, e como saída a composição da fase vapor do componente 1 ( $y_1$ ) e a temperatura ( $T$ ) do sistema.

Entretanto, após análise dessa proposta verificou-se que essa não iria funcionar, pois falta uma variável que identifique o comportamento de volatilidade entre os componentes do sistema binário. Analisando a Eq. (3.2) verificou-se que a variável que identifica essa relação é a pressão de saturação ( $P^{sat}$ ).

$$y_i P = g_i x_i P_i^{sat} \quad (3.2)$$

$P^{sat}$  é uma propriedade do componente puro e representa a sua volatilidade, portanto, adicionando-se  $P^{sat}$  dos dois componentes da mistura ter-se-ia uma relação das volatilidades do sistema e a RNA teria que simular apenas o valor de  $y_1$  e não mais o valor de  $x_1$  e  $P^{sat}$ . Portanto acrescentou-se uma nova entrada na rede, a variável  $P^{sat}_1/P^{sat}_2$ .

A rede R4 possui como entradas os grupos UNIFAC do componente 1, grupos UNIFAC do componente 2,  $P$ ,  $x_1$  e  $P^{sat}_1/P^{sat}_2$ , e como saídas  $y_1$  e  $T$ . A Figura 6.17 mostra a rede R4.

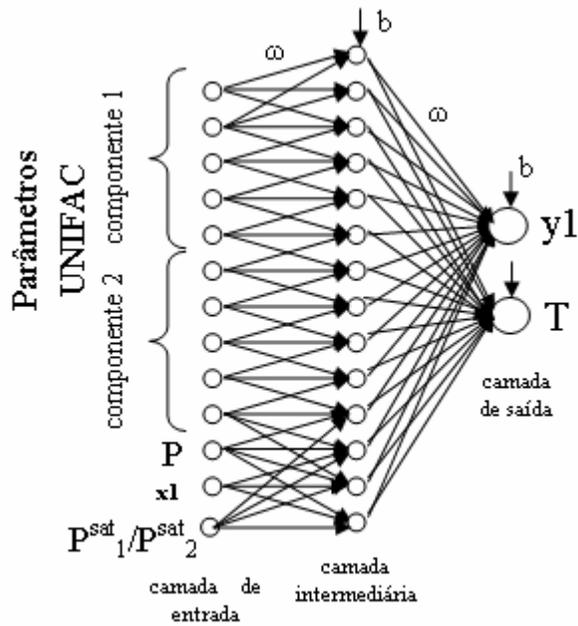


Figura 6.17 - Proposta 2, Rede R4.

Nessa rede foram usados os mesmos componentes e grupos da Tabela 6.1, ou seja, somente álcoois-água com os cinco grupos UNIFAC primeiramente estudados ( $\text{CH}_3$ ,  $\text{CH}_2$ ,  $\text{CH}$ ,  $\text{OH}$  e  $\text{H}_2\text{O}$ ). Foram usados como grupo de teste os dados etanol-água. A rede R4 possui 882 dados de treinamento e 198 dados de teste, com temperaturas de 15 a 215 °C e pressões de 1,29 a 5001 mmHg.

#### 6.4.1 Escolha do número de neurônios na camada escondida

A Tabela 6.10 apresenta o número de neurônios na camada escondida encontrados por três correlações para a rede R4.

Tabela 6.10 - Métodos utilizados para encontrar o número de neurônios na camada escondida para a rede R4.

<b>Correlações</b>	<b>Nº de Neurônios na Camada Escondida</b>
Hecht-Nielsen	<b>27</b>
Rogers e Jenkins	<b>111</b>
Dan Foresee e Hagan	<b>63</b>

Na tabela acima,  $J_1$  é igual a 13,  $J_1$  é igual a 2 e  $N_1$  é igual a 882.

A rede R4 foi simulada de 5 em 5 neurônios até o valor de 70 neurônios na camada escondida. Na faixa onde os valores começaram a se estabilizar a rede foi simulada variando o número de neurônios de uma em uma unidade (60, 61, 62, 63, 64 e 65). Cada neurônio foi simulado pelo menos três vezes, ao final dessa série de treinamento foi feita a média do SSE e do ?, essas médias foram os valores considerados nas análises e gráficos. Nas Figuras 6.18 e 6.19, o número de neurônios escolhido, no caso, 63 neurônios na camada intermediária, está representado em destaque por um triângulo ( $N = 1010,0$  e  $? = 575,8$ ).

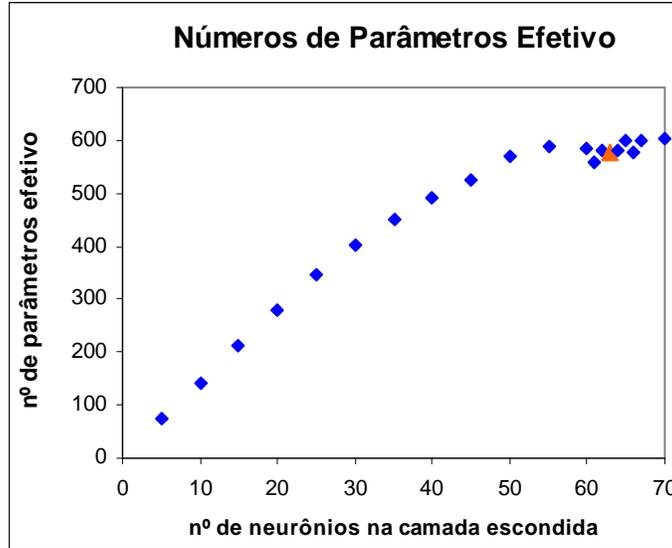


Figura 6.18 – Gráfico do ? pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 63.

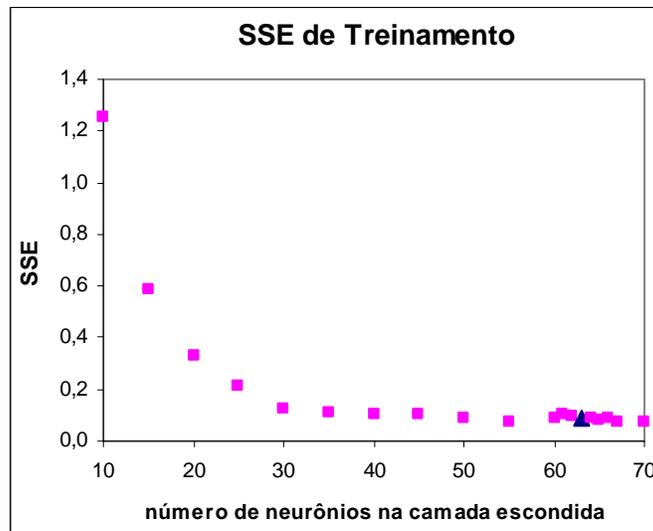


Figura 6.19 - Gráfico do SSE pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 63.

O ponto do neurônio 5 foi retirado da Figura 6.19 para melhor visualização do mesmo, já que seu valor divergia muito dos valores finais do SSE (para números de neurônios muito altos). O valor do SSE do neurônio 5 é igual a 6,33.

Observando-se a Tabela 6.10 nota-se que os valores fornecidos pelos três métodos são muito diferentes, sendo que, o valor do segundo método (Rogers e Jenkins) é novamente muito alto, sendo este inviável. O valor obtido pelo método de Dan Foresee e

Hagan também é considerado um valor alto e, como percebeu-se novamente, a rede R4 obteve um melhor desempenho com um menor número de neurônios na camada escondida, como é comparado e comprovado na próxima seção.

#### 6.4.2 Resultados da rede R4

A Tabela 6.11 e as Figuras 6.20 e 6.22 são referentes a 63 neurônios na camada escondida, a Tabela 6.12 e as Figuras 6.21 e 6.23 são referentes a 13 neurônios ( $N = 210,0$  e  $\eta = 186,45$ ). Os valores de  $y$  e  $T$  encontrados nas figuras abaixo são valores normalizados de 0 a 1.

Tabela 6.11 - Dados de treinamento e teste da rede R4 com 63 neurônios na camada escondida.

Neurônios de saída	y	T
SSE treinamento	0,0837	
R treinamento	1,000	
SSE teste	2,591	27.022
R teste	0,887	
MSE teste	0,0131	90,68

Tabela 6.12 - Dados de treinamento e teste da rede R4 com 13 neurônios na camada escondida.

Neurônios de saída	y	T
SSE treinamento	0,608	
R treinamento	0,999	
SSE teste	0,0578	2.359
R teste	0,997	
MSE teste	0,0003	11,91

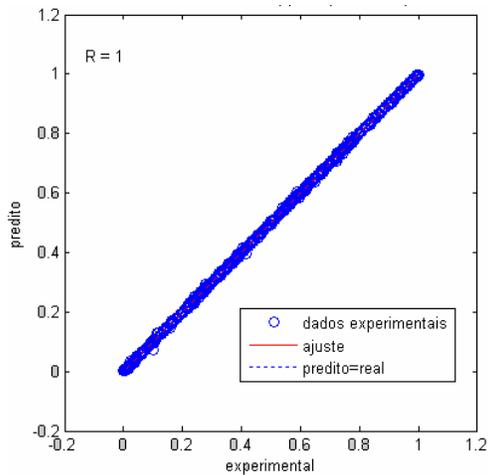


Figura 6.18 - Comparação entre valores experimentais de  $y$  e  $T$  com valores calculados pela rede neural R4 no treinamento, 63 neurônios na camada escondida.

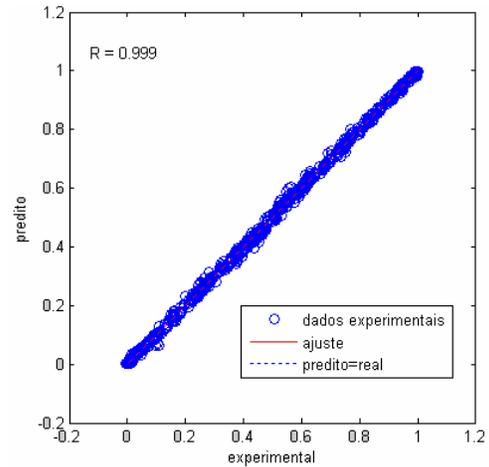


Figura 6.19 - Comparação entre valores experimentais de  $y$  e  $T$  com valores calculados pela rede neural R4 no treinamento, 13 neurônios na camada escondida.

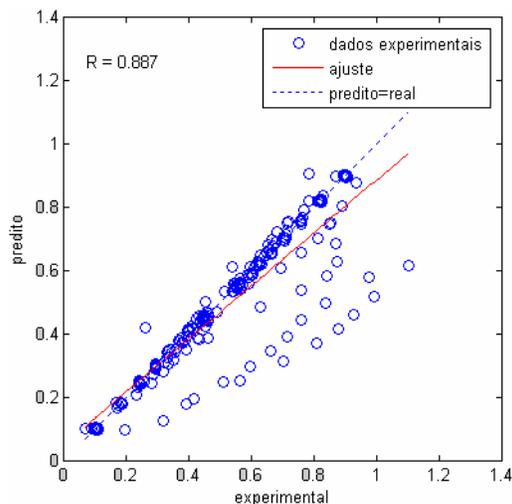


Figura 6.20 - Comparação entre valores experimentais de  $y$  e  $T$  com valores calculados pela rede neural R4 no teste, 63 neurônios na camada escondida.

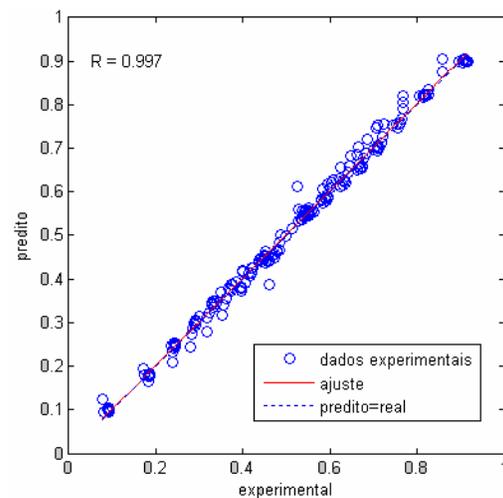


Figura 6.21 - Comparação entre valores experimentais de  $y$  e  $T$  com valores calculados pela rede neural R4 no teste, 13 neurônios na camada escondida.

Comparando-se as tabelas e gráficos referentes a 63 neurônios com as referentes a 13 neurônios percebe-se uma melhora do desempenho da rede para o caso em que a rede é simulada com um menor número de neurônios.

### 6.4.3 Conclusões parciais da rede R4

As Figuras 6.20 e 6.21 mostraram que o treinamento da rede apresentou um ótimo comportamento, tanto para 63 neurônios quanto para 13 neurônios, com os pontos ao longo da diagonal e R próximo de 1,0 nos dois casos.

Os resultados do gráfico referente ao teste da rede R4 para 13 neurônios mostraram uma grande melhora no desempenho da rede quando comparado com os resultados do gráfico referente a 63 neurônios. A Figura 6.23 apresentou seus pontos ao longo da diagonal com R perto de 1,0, ao contrário da Figura 6.22 que apresentou a reta de ajuste distante da diagonal com pontos dispersos pelo gráfico e R longe de 1,0. Esse resultado pode ser comprovado também através das Tabelas 6.11 e 6.12, nota-se uma diminuição do valor do MSE tanto para o  $y$  quanto para o  $T$ .

O desempenho da rede R4 foi considerado ótimo para 13 neurônios na camada intermediária. O MSE do  $y$  foi considerado desprezível (MSE = 0,0003). Já o MSE do  $T$  não foi tão baixo quanto o esperado, mas mesmo assim ainda é um valor viável para a faixa de temperaturas trabalhada, de 295,6 a 423,5 K, (MSE = 11,91).

## 6.5 Projeto da rede R5

A rede R5 faz a predição direta do ELV através do  $y$ . Essa rede tem as mesmas entradas e saídas da rede R4, mas os seus dados foram expandidos para os componentes e grupos usados na rede R3, Tabela 6.6. O grupo de teste também foi o mesmo usado na rede R3. A rede R5 possui 3240 dados de treinamento e 504 dados de teste, com temperaturas de 10 a 215 °C e pressões de 0,74 a 7109 mmHg.

### 6.5.1 Escolha do número de neurônios na camada escondida

Na Tabela 6.13 são apresentados o número de neurônios na camada escondida encontrados por três correlações para a rede R5.

Tabela 6.13 - Métodos utilizados para encontrar o número de neurônios na camada escondida para a rede R5.

Correlações	Nº de Neurônios na Camada Escondida
Hecht-Nielsen	<b>39</b>
Rogers e Jenkins	<b>295</b>
Dan Foresee e Hagan	<b>95</b>

Na tabela acima,  $J_1$  é igual a 19,  $J_2$  é igual a 2 e  $N_1$  é igual a 3240.

A rede R5 foi simulada de 5 em 5 neurônios até o valor de 100 neurônios na camada escondida. Cada neurônio foi simulado pelo menos três vezes, ao final dessa série de treinamento foi feita a média do SSE e do  $\phi$ , estas médias foram os valores considerados nas análises e gráficos. Nas Figuras 6.24 e 6.25, o número de neurônios escolhido, no caso, 95 neurônios na camada intermediária, está representado em destaque por um triângulo ( $N = 2092,0$  e  $\phi = 1401,2$ ).

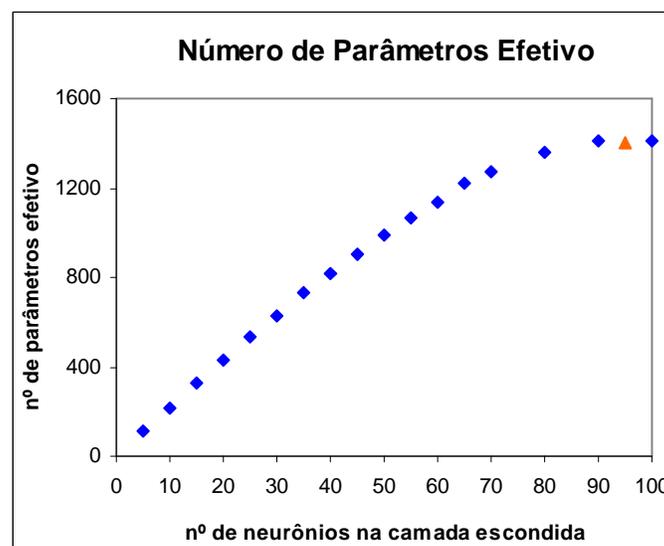


Figura 6.224 - Gráfico do  $\phi$  pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 63.

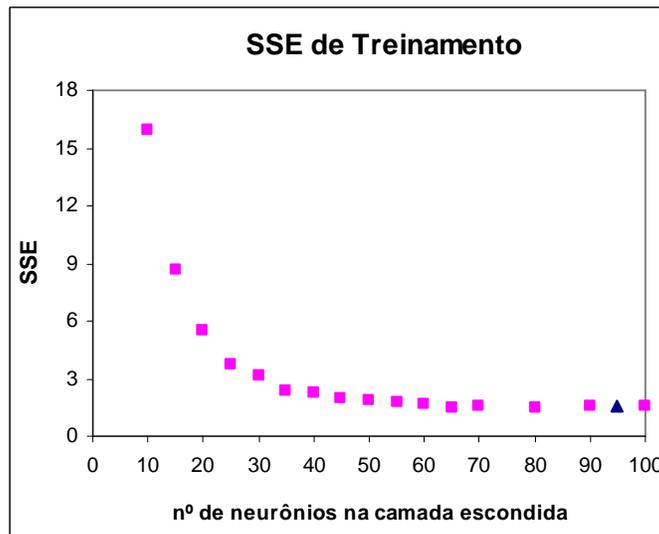


Figura 6.25 - Gráfico do SSE pelo número de neurônios na camada escondida. O triângulo no gráfico representa o número de neurônios escolhido, no caso neurônio número 95.

O ponto do neurônio 5 foi retirado da Figura 6.25 para melhor visualização do mesmo, já que seu valor divergia muito dos valores finais do SSE (para números de neurônios muito altos). O valor do SSE do neurônio 5 é igual a 55,05.

A Tabela 6.13 apresenta valores muito diferentes de cada método e, como nas redes anteriores, o segundo método fornece um valor mais uma vez inviável. O terceiro método também fornece um valor muito alto, mas como feito anteriormente, foram gerados os gráficos e tabelas para este número de neurônios (95) e para o número de neurônios da camada de entrada (19), para que seja possível a comparação do desempenho da rede.

### 6.5.2 Resultados da rede R5

A Tabela 6.14 e as Figuras 6.26 e 6.28 são referentes a 95 neurônios na camada escondida e a Tabela 6.15 e as Figuras 6.27 e 6.29 são referentes a 19 neurônios ( $N = 420,0$  e  $\tau = 407,48$ ). Os valores de  $y$  e  $T$  encontrados nas figuras abaixo são valores normalizados de 0 a 1.

Tabela 6.14 - Dados de treinamento e teste da rede R5 com 95 neurônios na camada escondida.

Neurônios de saída	y	T
SSE treinamento	1,551	
R treinamento	1,000	
SSE teste	6,038	159.944
R teste	0,936	
MSE teste	0,0120	317,35

Tabela 6.15 - Dados de treinamento e teste da rede R5 com 19 neurônios na camada escondida.

Neurônios de saída	y	T
SSE treinamento	5,190	
R treinamento	0,998	
SSE teste	1,640	161.024
R teste	0,981	
MSE teste	0,0033	319,49

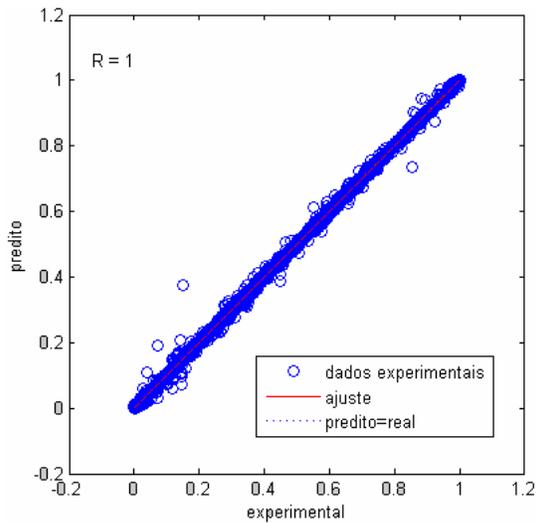


Figura 6.23 - Comparação entre valores experimentais de  $y$  e  $T$  com valores calculados pela rede neural R5 no treinamento, 95 neurônios na camada escondida.

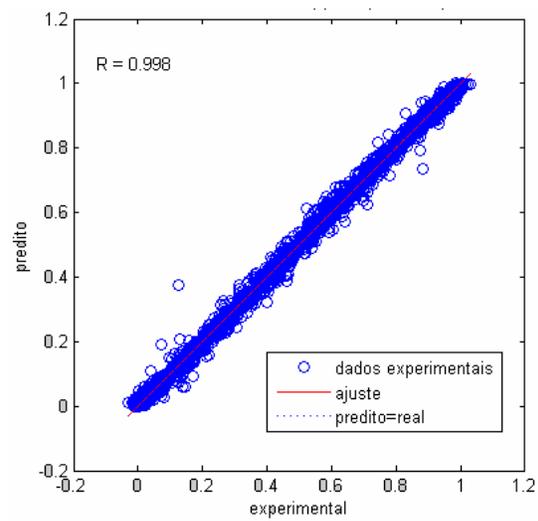


Figura 6.24 - Comparação entre valores experimentais de  $y$  e  $T$  com valores calculados pela rede neural R5 no treinamento, 19 neurônios na camada escondida.

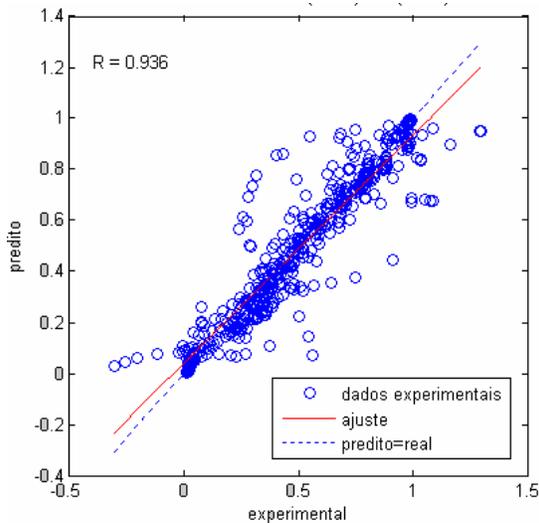


Figura 6.25 - Comparação entre valores experimentais de  $y$  e  $T$  com valores calculados pela rede neural R5 no teste, 95 neurônios na camada escondida.

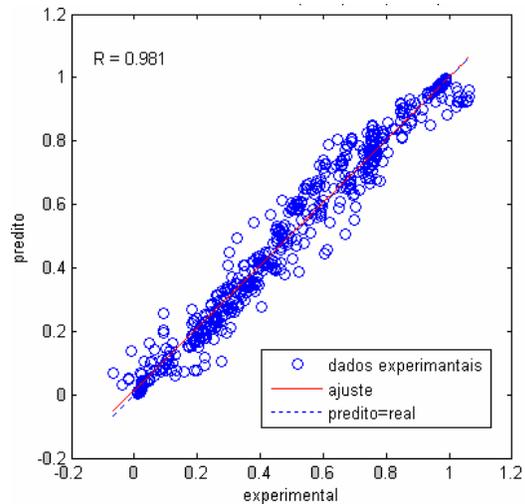


Figura 6.26 - Comparação entre valores experimentais de  $y$  e  $T$  com valores calculados pela rede neural R5 no teste, 19 neurônios na camada escondida.

Comparando as tabelas e gráficos referentes a 95 neurônios com os referentes a 19 neurônios, observa-se mais uma vez uma melhora no desempenho da rede para o segundo caso, ou seja, para um número de neurônios menor.

### 6.5.3 Conclusões parciais da rede R5

As Figuras 6.26 e 6.27, referentes ao treinamento da rede R5 apresentam ambas um bom desempenho, com a reta de ajuste na diagonal e com pontos do longo da reta.

Os resultados dos gráficos referentes ao teste da rede, Figuras 6.28 e 6.29, apresentam uma grande diferença entre si. O gráfico referente a 95 neurônios possui pontos dispersos e a reta de ajuste não está ao longo da diagonal. O gráfico referente a 19 neurônios apresenta a reta de ajuste ao longo da diagonal e seus pontos estão próximos da reta de ajuste. Esses resultados são comprovados através das Tabelas 6.14 e 6.15, o SSE de teste e MSE do  $y$  diminuíram muito os seus valores, podendo-se considerar o valor do MSE do  $y$  para 19 neurônios desprezível ( $MSE = 0,0033$ ). Já em relação à variável  $T$ , o SSE de teste e MSE apresentaram em ambas as tabelas um valor muito alto. Esses valores não

podem ser ignorados já que o MSE para 19 neurônios apresenta um valor na mesma ordem de grandeza da faixa de temperatura trabalhada, de 278,76 a 485,87 (MSE = 319,49).

Outro problema encontrado na rede R5 é que o  $y$  apresentou em alguns momentos valores menores do que zero ou maiores do que um. Isso já era um problema esperado, visto que não se podem colocar restrições em uma RNA e as regiões perto dos extremos são geralmente as regiões de maior não idealidade de uma mistura.

A rede R5 foi considerada uma rede com desempenho razoável pois apresentou ótimos resultados para a variável  $y$  e resultados ruins para a variável  $T$ . Mas é de se considerar a complexidade dessa rede que, como a rede R3, possui muitos tipos de componentes com comportamentos muito não-lineares.

## 7 Conclusão

O objetivo dessa dissertação foi verificar se as redes neurais artificiais (RNAs) eram capazes de prever o equilíbrio líquido-vapor (ELV) de misturas binárias não ideais utilizando a abordagem do método de contribuição de grupos UNIFAC. Para isso foram feitas duas propostas de rede, uma com as RNAs substituindo o modelo UNIFAC no cálculo do coeficiente de atividade ( $\gamma$ ) e outra calculando diretamente a composição da fase vapor ( $y$ ) através das RNAs. Para os dois casos foram usados, inicialmente somente um tipo de mistura binária, álcoois-água com cinco grupos funcionais UNIFAC. Após a verificação do bom desempenho da rede, novos componentes e grupos funcionais UNIFAC foram acrescentados (aldeídos, cetonas e mais três grupos funcionais UNIFAC).

As redes referentes à primeira proposta foram consideradas boas, especialmente no caso da rede que trabalhava somente com componentes álcoois-água. Mesmo assim, nos dois casos houve pontos com grandes erros, principalmente na faixa perto dos extremos da curva de ELV ( $x \rightarrow 0$  e  $x \rightarrow 1$ ), pois essa região apresenta normalmente uma grande não idealidade.

As redes referentes à segunda proposta apresentaram desempenhos diferenciados. No caso da rede que usou somente os componentes álcoois-água, a rede apresentou ótimos resultados, tanto para a variável  $y$  quanto para a variável  $T$ . Já a rede que utilizou mais componentes e grupos UNIFAC obteve um ótimo desempenho para a variável  $y$  e um desempenho ruim para a variável  $T$ , pois os erros para essa variável foram da mesma ordem de grandeza que a faixa de temperatura trabalhada. Para as redes que utilizam a segunda proposta foi notado que em alguns casos a variável  $y$  apresentou valores menores do que 0 e maiores do que 1, o que é termodinamicamente infactível. Esse problema já era esperado, visto que não se podem colocar restrições em uma RNA e as regiões perto dos extremos são geralmente as locais de maior não idealidade de uma mistura.

Os resultados das redes foram considerados bons, em geral. Todas as redes apresentaram uma boa capacidade de predição, com desvios na mesma ordem de grandeza e algumas vezes até menores do que os métodos baseados no uso de um modelo específico para a descrição do comportamento ELV, como o método UNIFAC, por exemplo. Mas sem

a necessidade de ser definido um modelo ou de serem calculados parâmetros de interação ou constantes ajustáveis. Este resultado é relevante, uma vez que os sistemas estudados apresentam não idealidades acentuadas.

Quanto ao número de neurônios na camada escondida foram mostrados três métodos de cálculo do mesmo e escolhido o terceiro método, Dan Foresee e Hagan (1997), como o método utilizado nas simulações, sendo que os outros foram mostrados apenas a título de comparação. O método de Dan Foresee e Hagan foi o escolhido por apresentar um estudo mais aprofundado das RNAs estudadas, pois este depende de parâmetros da rede simulada (SSE,  $N$  e  $?$ ) e não de parâmetros gerais, como: o número de neurônios na camada de entrada e saída e o número de dados de treinamento. Mesmo assim, os melhores resultados foram obtidos pelas redes com números de neurônios menores do que os indicados pelos três métodos. Esse resultado demonstra a independência de teorias e postulados na área de redes neurais artificiais, já que não existe como prever o seu desempenho e nem como escolher através de uma forma ótima e única os seus algoritmos e funções.

## **8 Sugestões para trabalhos futuros**

O presente trabalho pode ser continuado através de pesquisas como:

- Melhorar a base de dados de entrada da rede, colocando mais pontos de cada componente em toda a faixa de pressão e temperatura a ser estudada;
- Otimizar as redes, trabalhando com diversos algoritmos de treinamento, funções de ativação e outros parâmetros das RNAs.

## 9 Referências Bibliográficas

DAN FORESEE, F.; HAGAN, M.T. Gauss-Newton approximation to Bayesian learning, *In The International Joint Conference on Neural Networks*, p. 1930-1935. 1997.

DE CASTRO, L.N. *Fundamentals of natural computing. Basic concepts, algorithms, and applications*. Chapman & Hall/ CRC. 2006.

DECHEMA. GMEHLING, J.; ONKEN, U.; WEIDLICH, U. *Vapor-liquid equilibrium data collection. DECHEMA, Chemistry Data Series. Vol.I, Part.1. Vol. I., Part. 1<sup>a</sup>. Vol. I, Part. 2c. Vol. I, Part. 2d. Vol. I, Part. 2f. Vol. I, Part 3a. Vol. I, Part. 3b*. 1982.

FREDENSLUND, A.; GMEHLING, J.; RASMUSSEN, P. *Vapor-liquid equilibria using UNIFAC a group-contribution method*. Elsevier Scientific Publishing Company. 1977.

GANGULY, S. *Prediction of VLE data using radial basis function network*. *Computers and Chemical Engineering*, 27, 1445-1454. 2003.

HAGAN, M.T.; DEMUTH, H.B.; BEALE, M. *Neural network design*. PWS Publishing Company. 1996.

HAYKIN, S. *Neural networks, a comprehensive foundation*. Macmillan College Publishing Company. 1994.

ILIUTA, M.C.; ILIUTA, I.; LARACHI, F. *Vapour-liquid equilibrium data analysis for mixed solvent-electrolyte systems using neural network models*. *Chemical Engineering Science*, 55, pages 2813-2825. 2000.

MOHANTY, S. *Estimation of vapour liquid equilibria for the system carbon dioxide-difluormethane using artificial neural networks*. International Journal of Refrigeration, 29, pages 243-249. 2006.

MOHANTY, S. *Estimation of vapour liquid equilibria of binary systems, carbon dioxide-ethyl caproate, ethyl caprylate and ethyl caprate using artificial neural networks*. Fluid Phase Equilibria, 235, pages 92-98. 2005.

NGUYEN, V.D.; TAN, R.R.; BRONDIAL, Y.; FUCHINO, T. *Prediction of vapor-liquid equilibrium data for ternary systems using artificial neural networks*. Fluid Phase Equilibria, 254, pages 188-197. 2007.

PASTI, R. *Redes Neuro-Imunes Aplicadas a Classificação e Otimização Combinatória*. Ciência da Computação, Universidade Católica de Santos. Tese (Mestrado). 2007.

PETERSEN, R.; FREDENSLUND, A.; RASMUSSEN, P. *Artificial neural networks as a predictive tool for vapor-liquid equilibrium*. Computers and Chemical Engineering, vol. 18, Suppl., pages S63-S67. 1994.

PIOTROWSKI, K.; PIOTROWSKI, J.; SCHLESINGER, J. *Modelling of complex liquid-vapour equilibria in the urea synthesis process with the use of artificial neural network*. Chemical Engineering and Processing, vol. 42, issue 4, pages 285-289. 2003.

REID, R.C.; PRAUSNITZ, J.M.; POLING, B.E. *The properties of gases and liquids*. 4th Edition, McGraw-Hill Book Company. 1987.

SANDLER, S.I. *Chemical and engineering thermodynamics*. John Wiley & Sons Published. 1989.

SEN, M.; YANG, K.T. Applications of artificial neural networks and genetic algorithms in thermal engineering, em *CRC Handbook, Thermal Engineering*. (editor: Kreith, F.). 2000.

SHARMA, R.; SINGHAL, D.; GHOSH, R.; DWIVEDI, A. *Potencial applications of artificial neural networks to thermodynamics: vapor-liquid equilibrium predictions*. Computers and Chemical Engineering, 23, pages 385-390. 1999.

SCHMITZ, J.E. *Modelos híbridos de colunas de destilação*. Faculdade de Engenharia Química, Universidade Estadual de Campinas. Tese (Mestrado). 2002.

SCHMITZ, J.E. *Cálculos de estabilidade e divisão de fases por meio de redes neurais artificiais*. Faculdade de Engenharia Química, Universidade Estadual de Campinas. Tese (Doutorado). 2006.

SMITH, J.M.; VAN NESS, H.C.; ABBOTT, M.M. *Introdução á termodinâmica da engenharia química*. LTC – Livros Técnicos e Científicos Editora S.A., 5ª Edição. 2000.

URATA, S; TAKADA, A.; MURATA, J.; HIAKI, T.; SEKIYA, A. *Prediction of vapor-liquid equilibrium for binary systems containing HFEs by using artificial neural network*. Fluid Phase Equilibria, 199, pages 63-78. 2002.

WALAS, S.M. *Phase equilibria in chemical engineering*. Butterworth Publishers. 1985.