

UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA QUÍMICA

ÁREA DE CONCENTRAÇÃO
SISTEMAS DE PROCESSOS QUÍMICOS E INFORMÁTICA

MODELOS HÍBRIDOS DE COLUNAS DE DESTILAÇÃO

Autor: Jones Erni Schmitz

Orientador: Mário de Jesus Mendes

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

Campinas – São Paulo

Abril de 2002

UNICAMP
BIBLIOTECA CENTRAL

UNIDADE BAE-FE 12
Nº CHAMADA T/UNICAMP
Sch56m
V _____ EX _____
TOMBO BCI 49436
PROC 16.83710 @
C _____ D Y _____
PREÇO R\$ 11,00
DATA _____
Nº CPD _____

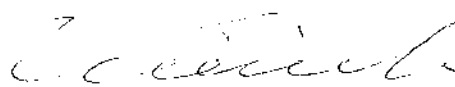
CM00168404-1

313 ID 242784

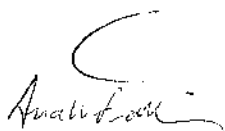
FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

Sch56m Schmitz, Jones Erni
Modelos híbridos de colunas de destilação / Jones Erni Schmitz.--Campinas, SP: [s.n.], 2002.
Orientador: Mário de Jesus Mendes.
Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Química.
1. Métodos de simulação. 2. Redes neurais (Computação). 3. Modelos matemáticos. I. Mendes, Mário de Jesus. II. Universidade Estadual de Campinas. Faculdade de Engenharia Química. III. Título.

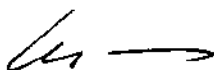
Dissertação de Mestrado defendida por Jones Erni Schmitz e aprovada em 16 de
Abril de 2002 pela banca examinadora constituída pelos doutores:



Prof. Dr. Mário de Jesus Mendes – Orientador

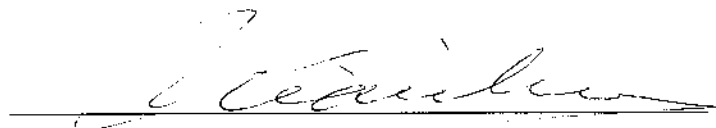


Prof.ª Dr.ª Ana Maria Frattini Fileti



Prof. Dr. Roger Josef Zemp

Esta versão corresponde à versão final Dissertação de Mestrado em Engenharia
Química.

A handwritten signature in black ink, written in a cursive style, positioned above a horizontal line. The signature appears to read 'Mário de Jesus Mendes'.

Prof. Dr. Mário de Jesus Mendes - Orientador

Agradecimentos

Ao Prof. Dr. Mário de Jesus Mendes pela orientação, paciência e atenção dispensadas.

A CAPES, pelo apoio financeiro.

Ao Prof. Dr. Roger Josef Zemp por ceder seu pacote de cálculos termodinâmicos.

À minha família.

A todos os funcionários da FEQ, em especial à Andréia e Joseane.

Ao colega Alexandre Rodrigues Simões pelas dicas e recomendações.

A meus colegas de mestrado e de república.

Resumo

O emprego de técnicas on-line de controle preditivo por modelo e de otimização de processos requer que a simulação destes processos seja muito rápida. Porém em alguns casos os modelos dos processos são muito complexos e de elevada dimensão. Nestes casos para uma simulação suficientemente rápida devem ser empregadas técnicas de aceleração. Entre estas, estão incluídas as técnicas de redução de ordem e de simplificação do modelo.

Um exemplo de processo onde estes problemas são encontrados é o das colunas de destilação. Neste caso, além das técnicas de redução de ordem baseadas no emprego de colocação ortogonal discreta, podem ser usadas técnicas de simplificação de modelos baseados no fato de que 30 a 80% do tempo gasto na simulação são usados no cálculo das propriedades termodinâmicas.

Neste contexto surge a possibilidade do emprego de redes neurais no cálculo das propriedades termodinâmicas. Uma vez que estas podem representar muito bem, modelos não-lineares, como é o caso das propriedades de equilíbrio líquido-vapor, e realizar os cálculos muito rapidamente. Deste modo, o modelo da coluna de destilação torna-se um modelo híbrido, formado pelas equações de balanço de massa e energia nos pratos e pela rede neural construída para predizer as propriedades termodinâmicas.

O emprego deste modelo híbrido na simulação estática de uma coluna de destilação implicou na redução do tempo computacional em cerca de 55%. Demonstrando, desta forma, a viabilidade do emprego de redes neurais na predição de dados termodinâmicos para a simulação de colunas de destilação.

Palavras Chave: Otimização, Redes Neurais Artificiais, Colunas de Destilação.

Abstract

Techniques like online process optimization and model predictive control require fast simulations of the process models used. However, in some cases the process models are very complex or have a high dimension. In these cases acceleration techniques that include order reduction and model simplification can be used to obtain fast enough simulations.

Distillation columns are an example of process where this type of problem is usually found. In this case, besides order reduction techniques based on discrete orthogonal collocation, simplification techniques have been developed, like local thermodynamic models, based on the fact that 30-80 % of the simulation time were associated with thermodynamic calculations.

In this context, the possibility arose of using neural networks to predict thermodynamic properties. It is well known that neural networks are able to represent non-linearities like those found in thermodynamic models. Neural network models can be combined with physical models, based on mass and energy balances, for example, to obtain the so called hybrid models.

In this work a static hybrid model of a binary (propane/propene) distillation column has been developed, using neural networks for the calculation of equilibrium constants, of enthalpies and of their derivatives used in the simulation. When used for the simulation of the column, this model allowed for a reduction of about 55 % of the computational time, in comparison with a classical model where the thermodynamic properties were obtained using a method based on the Peng-Robinson state equation.

Keywords: Optimization, artificial neural networks, distillation columns.

Sumário

Agradecimentos	iii
Resumo	iv
Abstract	v
Nomenclatura	ix
1. Introdução	1
2. Análise de Colunas de Destilação de Pratos em Regime estacionário	3
2.1. Modelagem de Colunas de Destilação de Pratos em Regime Estacionário	3
2.1.1. Modelo de um Prato Genérico	4
2.1.2. Modelo do Refervedor de Equilíbrio	6
2.1.3. Modelos para o Condensador	6
2.1.4. Análise de Graus de Liberdade	10
2.2. Simulação de Colunas de Destilação de Pratos	11
2.2.1. Método de Naphtali-Sandholm	12
2.2.2. Cálculo das Propriedades Termodinâmicas	14
2.2.3. Especificações da Coluna	16
2.2.4. Resultados da Simulação	17
2.3. Conclusões	23
3. Emprego de Redes Neurais Artificiais na Predição de Propriedades Termodinâmicas	24
3.1. Introdução	24
3.2. Fundamentos de Redes Neurais Artificiais	25

3.2.1. Modelo de um Neurônio	25
3.2.2. Arquitetura de Redes Neurais Artificiais	27
3.2.3. Utilização de Redes Neurais Artificiais	30
3.2.4. Vantagens e Limitações de Redes Neurais Artificiais	30
3.3. Técnicas de Treinamento	31
3.3.1. Algoritmo de Levenberg-Marquardt	33
3.4. Melhorando a Generalização	34
3.4.1. Regularização	34
3.4.2. Otimização dos Parâmetros de Regularização pela Lei de Bayes – Regularização Bayesiana	35
3.5. Redes Neurais Artificiais na Predição de Propriedades Termo- dinâmicas	38
3.5.1. Normalização dos Padrões de Entrada-Saída	39
3.5.2. Resultados Obtidos	40
3.6. Conclusões	45
4. Modelos Estáticos Híbridos de Colunas de Destilação	46
4.1. Introdução	46
4.2. Modelos Híbridos	47
4.3. Metodologia Empregada	47
4.4. Resultados	50
4.4.1. Condensador Parcial	50
4.4.2. Condensador Total	52
4.5. Conclusões	53
5. Discussão dos Resultados	54
5.1. Sugestões para Trabalhos Futuros	54

Referências Bibliográficas	55
Apêndice A – Método de Naphtali-Sandholm	59
Apêndice B – Equação de Peng-Robinson	60

Nomenclatura

Simbologia

<i>D</i>	vazão de destilado	Mmol.dia ⁻¹
<i>E</i>	somatório quadrático dos erros	
<i>F</i>	vazão total da corrente de alimentação	Mmol.dia ⁻¹
<i>K</i>	constante de equilíbrio	
<i>l</i>	vazão molar do componente na fase líquida	Mmol.dia ⁻¹
<i>L</i>	vazão molar total de líquido	Mmol.dia ⁻¹
<i>N</i>	número de parâmetros da rede neural	
<i>n</i>	número de padrões usado no treinamento de redes neurais	
<i>nf</i>	prato da alimentação	
<i>H</i>	entalpia total da corrente	kJ.dia ⁻¹
<i>h</i>	entalpia molar do componente na corrente	kJ.kmol ⁻¹
<i>P</i>	pressão	kPa
<i>Q</i>	calor	kJ.dia ⁻¹
<i>R</i>	razão de refluxo	
<i>T</i>	temperatura	°C
<i>v</i>	vazão molar do componente na fase vapor	kmol.dia ⁻¹
<i>V</i>	vazão molar total de vapor	kmol.dia ⁻¹
<i>x</i>	fração molar do componente na fase líquida	
<i>y</i>	fração molar do componente na fase vapor	
<i>z</i>	fração molar do componente na alimentação	

Letras Gregas

α	parâmetro de regularização
β	parâmetro de regularização
γ	número efetivo de parâmetros

Índice Inferior

nst	número total de estágios
nc	número de componentes
W	pesos sinápticos
D	conjunto de treinamento das redes neurais artificiais

Índice Superior

L	fase líquida
V	fase vapor

1 – Introdução

O emprego de colunas de destilação como método de separação é muito difundido na Indústria Química, encontrando aplicação nas mais diversas áreas, abrangendo desde indústria química pesada até indústria alimentícia. Devido a esta importância e às peculiaridades do controle, modelagem, simulação e projeto, as colunas de destilação têm sido amplamente estudadas. Uma das principais dificuldades enfrentadas nestes estudos está relacionada à elevada dimensão e à complexidade das equações do modelo resultante, as quais necessitam de grandes esforços computacionais para sua resolução. Como resultado, grande parte destes estudos, têm buscado técnicas que levam a redução da carga computacional por meio de redução da ordem ou simplificação das equações do modelo.

A solução destes modelos reduzidos ou simplificados deve convergir para a do modelo rigoroso e deve poder ser obtida no menor tempo possível. Modelos com estas características são desejáveis quando se pretende utilizar técnicas de controle moderno e de otimização em tempo real.

Na otimização de processos em tempo real deseja-se conhecer as condições de operação ótimas deste processo. Para tanto, se deve usar um modelo que representa o comportamento do processo no momento, sem exigir muito esforço computacional. No caso de colunas de destilação, o emprego de modelos reduzidos e/ou simplificados se faz necessário uma vez que a simulação de um modelo completo exige muito tempo computacional. Já o emprego de algumas técnicas de controle moderno, tal como o controle preditivo com modelo requer a simulação dinâmica do modelo para calcular a trajetória da variável controlada a cada instante de amostragem. Neste caso, o tempo requerido para a otimização e simulação do modelo deve ser reduzido, pois quanto menor o tempo necessário para se obter uma ação de controle, menor pode ser o intervalo de amostragem, de maneira que se podem obter produtos dentro de especificações mais rígidas.

O objetivo deste trabalho é testar o emprego de um modelo híbrido na simulação de uma coluna de destilação. Este modelo empregará redes neurais artificiais para a predição das propriedades termodinâmicas, juntamente com as equações resultantes da modelagem matemática da coluna de destilação.

Este trabalho foi dividido em três partes: o Capítulo 2 tratará sobre a análise de colunas de destilação; no Capítulo 3 o emprego de redes neurais na predição de propriedades termodinâmicas será testado; e no Capítulo 4 o modelo híbrido será empregado na simulação de uma coluna de destilação.

2 – Análise de Colunas de Destilação de Pratos em Regime Estacionário

Neste capítulo tratar-se-á da análise de colunas de destilação de pratos em regime estacionário. Inicialmente será feita a modelagem deste tipo de coluna. Então será feita a simulação, usando o método de Naphtali e Sandholm (1971). Por fim, serão apresentados os resultados obtidos na simulação de uma determinada coluna de destilação.

2.1 - Modelagem de Colunas de Destilação de Pratos em Regime Estacionário

A destilação é um processo interfases, onde há pelo menos uma fase líquida em contato com uma fase vapor em cada estágio. Se o tempo de contato entre as fases em cada estágio for suficiente, as espécies químicas distribuem-se entre as fases de acordo com os critérios de equilíbrio termodinâmico. Após a separação das fases pode-se dizer que estas alcançaram o equilíbrio.

Esta é a principal suposição do modelo de estágio em equilíbrio, que é o mais usado na simulação de colunas de destilação de pratos. As equações que compõem o modelo de estágio de equilíbrio são conhecidas como equações MESH (Wang e Henke, 1966).

Para a modelagem da coluna de destilação serão feitas as seguintes hipóteses:

- 1) Retenção de vapor nos estágios desprezível;
- 2) Pressão constante ao longo da coluna (P);
- 3) Refervedor e pratos tratados como estágios de equilíbrio;

A modelagem da coluna de destilação de pratos será dividida em três partes: a modelagem de um prato genérico, a modelagem do refervedor de equilíbrio e a modelagem do condensador. Esta divisão é devida as características próprias das equações de cada um destes estágios. A coluna de destilação a ser modelada, é uma coluna simples com uma só

corrente de alimentação e com retiradas apenas no topo (condensador) e no fundo (refervedor).

2.1.1 - Modelo de um Prato Genérico

De acordo com a hipótese do estágio de equilíbrio, as equações do modelo para um prato genérico da coluna, como o ilustrado na Figura 2.1, são:

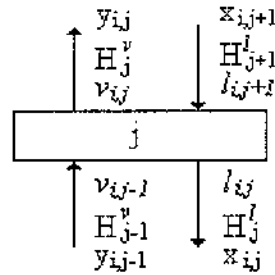


Figura 2.1 - Prato genérico.

Equações de balanço de massa por componente:

$$M_{i,j} \equiv l_{i,j+1} - l_{i,j} + v_{i,j-1} - v_{i,j} = 0, \quad \begin{matrix} i = 1, \dots, nc \\ j = 2, \dots, nst - 1 (j \neq nf) \end{matrix} \quad (2.1)$$

onde l_{ij} é a vazão do componente i na corrente líquida que sai do prato j , v_{ij} é a vazão do componente i na corrente de vapor que sai do prato j , nc é o número total de componentes e nst é o número total de pratos da coluna (incluindo o condensador e o fervedor).

Equações de balanço de energia:

$$H_j \equiv H_{j+1}^l - H_j^l + H_{j-1}^v - H_j^v = 0, \quad j = 2, \dots, nst - 1 (j \neq nf) \quad (2.2)$$

onde H_j^l é a entalpia da corrente líquida do prato j e H_j^v é a entalpia da corrente de vapor do prato j .

Relações de equilíbrio:

$$E_{i,j} \equiv K_{i,j} \frac{V_j}{L_j} - v_{i,j} = 0, \quad \begin{array}{l} i = 1, \dots, nc \\ j = 1, \dots, nst \end{array} \quad (2.3)$$

onde $K_{i,j}$ é a constante de equilíbrio líquido-vapor do componente i no prato j . Por definição tem-se ainda:

$$L_j = \sum_{i=1}^{nc} l_{i,j}, \quad (2.4)$$

$$V_j = \sum_{i=1}^{nc} v_{i,j}, \quad (2.5)$$

$$y_{i,j} = \frac{v_{i,j}}{V_j}, \quad (2.6)$$

$$x_{i,j} = \frac{l_{i,j}}{L_j}, \quad (2.7)$$

$$H_j^l = \sum_{i=1}^{nc} l_{i,j} h_{i,j}^l, \quad (2.8)$$

$$H_j^v = \sum_{i=1}^{nc} v_{i,j} h_{i,j}^v, \quad (2.9)$$

L_j é a vazão total de líquido do prato j , V_j é a vazão total de vapor do prato j , $x_{i,j}$ é a fração molar do componente i no líquido do prato j , $y_{i,j}$ é a fração molar do componente i no vapor do prato j , $h_{i,j}^l$ é a entalpia parcial molar do componente i na corrente líquida do prato j e $h_{i,j}^v$ é a entalpia parcial molar do componente i na corrente de vapor do prato j .

Para o prato da alimentação é necessário levar em consideração as vazões mássicas e de energia na corrente de alimentação.

2.1.2 - Modelo do Refervedor de Equilíbrio

As equações para um refervedor como o mostrado na Figura 2.2 são apresentadas a seguir:

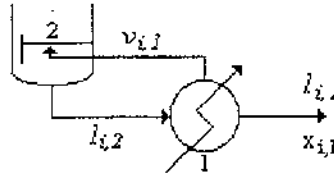


Figura 2.2 - Refervedor de equilíbrio.

Equações de balanço de massa por componente:

$$M_{i,1} \equiv l_{i,2} - l_{i,1} - v_{i,1} = 0, \quad i = 1, \dots, nc \quad (2.10)$$

Equações de balanço de energia:

$$H_1 \equiv H_2^l - H_1^l - H_1^v + Q_1 = 0, \quad (2.11)$$

onde Q_1 é a quantidade de calor adicionada ao refervedor.

Relações de equilíbrio:

$$E_{i,1} \equiv K_{i,1} l_{i,1} \frac{V_1}{L_1} - v_{i,1} = 0, \quad i = 1, \dots, nc. \quad (2.12)$$

2.1.3 - Modelos para o Condensador

Dois casos podem ser considerados: condensador parcial e condensador total.

a. Condensador Parcial

Um condensador parcial pode operar de duas maneiras, tal como mostra a Figura 2.3.

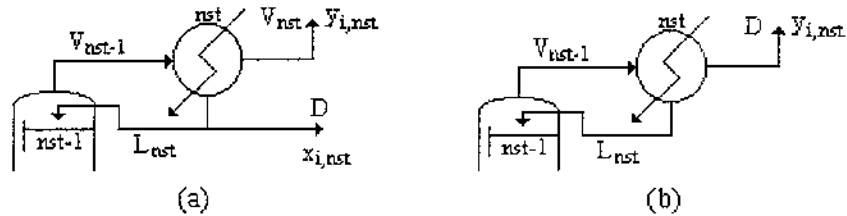


Figura 2.3 - Esquemas de funcionamento de condensadores parciais: (a) Condensador parcial com retiradas na fase líquida e vapor; (b) Condensador parcial com retirada somente da fase vapor.

A Figura 2.4 apresenta o esquema de uma coluna de destilação operando com um condensador parcial do tipo mostrado na Figura 2.3.b.

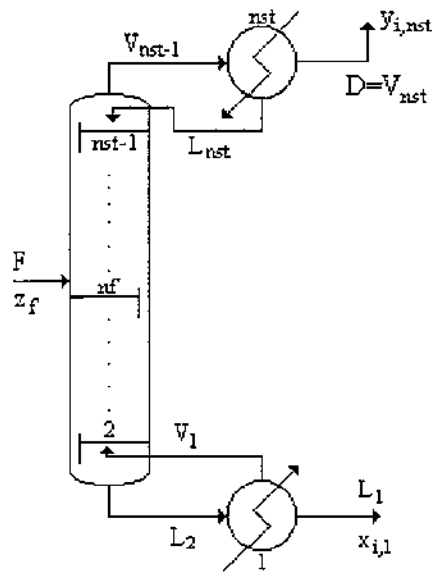


Figura 2.4 – Coluna de destilação com condensador parcial.

Neste caso a razão de refluxo, R , é definida por:

$$R = \frac{D}{L_{nst}}$$

onde D é a vazão de destilado.

Equações de balanço de massa por componente no condensador:

$$M_{i,nst} \equiv v_{i,nst-1} - l_{i,nst} - v_{i,nst} - Dx_{i,nst} = 0, \quad i = 1, \dots, nc \quad (2.13)$$

Equações de balanço de energia no condensador:

$$H_{i,nst} \equiv H_{nst-1}^v + H_{nst}^l - H_{nst}^v - H_D - Q_{nst} = 0, \quad (2.14)$$

onde H_D é a quantidade de energia contida na corrente de destilado e Q_{nst} é a quantidade de calor retirada do condensador.

Relações de equilíbrio no condensador:

$$E_{i,nst} \equiv K_{i,nst} l_{i,nst} \frac{V_{nst}}{L_{nst}} - v_{i,nst} = 0 \quad i = 1, \dots, nc \quad (2.15)$$

b. Condensador Total

Um condensador total não pode ser considerado estágio de equilíbrio. O esquema de uma coluna de destilação operando com um condensador total é ilustrado na Figura 2.5.

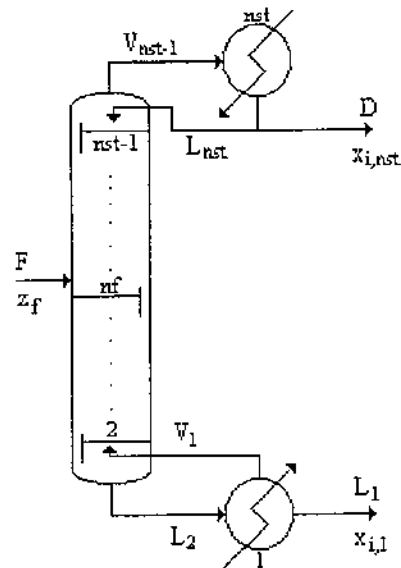


Figura 2.5 – Coluna de destilação com condensador total.

Equações de balanço de massa por componente no condensador:

$$M_{i,nst} \equiv v_{i,nst-1} - l_{i,nst} - Dx_{i,nst} = 0, \quad i = 1, \dots, nc \quad (2.16)$$

Equações de balanço de energia no condensador:

$$H_{i,nst} \equiv H_{nst-1}^v + H_{nst}^l - H_D - Q_{nst} = 0, \quad (2.17)$$

Critério de temperatura:

$$T_{nst} - T_{PB} - \Delta T = 0 \quad (2.18)$$

onde T_j é temperatura da saída do condensador e T_{PB} é a temperatura de ponto de bolha.

2.1.4 - Análise de graus de liberdade

O número de graus de liberdade é obtido subtraindo-se o número de equações do modelo do número de variáveis. Um grau de liberdade é uma variável cujo valor é fixado para definir o estado de operação da coluna.

a. Condensador Parcial

Número de equações do modelo:

i) por estágio

Balanços de massa por componente	<i>nc</i>
Balanços de energia	1
Relações de equilíbrio de fases	<i>nc</i>
Total	2.nc+1

ii) para a coluna inteira:

Total	nst.(2.nc+1)
--------------	---------------------

Número de variáveis do modelo (coluna da Figura 2.4):

Vazões molares dos componentes na fase líquida	<i>nst.nc</i>
Vazões molares dos componentes na fase vapor	<i>nst.nc</i>
Temperatura	<i>nst</i>
Pressão	1
Calor cedido ao refeedor	1
Calor retirado do condensador	1
Total	nst.(2.nc+1)+3

Graus de liberdade:

$$G.L. = [nst.(2.nc+1)+3] - [nst.(2.nc+1)]$$

$$G.L. = 3$$

Então, para a resolução das equações do modelo, é necessário especificar o valor de três variáveis. Admitindo-se que a pressão na coluna é dada, então as outras duas variáveis podem ser, por exemplo; a vazão de destilado e a razão de refluxo. As variáveis calculadas são as frações molares em cada fase, as temperaturas em cada estágio, as vazões molares em ambas as fases e as quantidades de calor transferidas no condensador e no refeedor.

b. Condensador Total

Número de equações do modelo:

i) por prato genérico incluindo refeedor:	
Balanços de massa por componente	<i>nc</i>
Balanços de energia	1
Relações de equilíbrio de fases	<i>nc</i>
ii) para o condensador:	
Balanços de massa por componente	<i>nc</i>
Balanços de energia	1
Equação (1.18)	1
iii) para a coluna inteira:	
Total	<i>nst.(2.nc+1)-nc+1</i>

Número de variáveis do modelo:

Vazões molares dos componentes na fase líquida	<i>nst.nc</i>
Vazões molares dos componentes na fase vapor	<i>(nst-1).nc</i>
Temperatura	<i>nst</i>
Pressão	1
Vazão de destilado	1
Calor cedido ao refeedor	1
Calor retirado do condensador	1
Total	<i>nst.(2.nc+1)-nc+4</i>

Graus de liberdade:

$$G.L. = [nst.(2.nc+1)-nc+4] - [nst.(2.nc+1)-nc+1]$$

$$G.L. = 3$$

Como se vê, o número de graus de liberdade da coluna não é alterada pela alteração do tipo de condensador.

2.2 – Simulação de Colunas de Destilação de Pratos

Nas últimas décadas vários métodos para simulação de colunas de destilação, ou seja para a solução das equações do modelo da coluna, têm sido propostos.

Estes métodos podem ser basicamente divididos em duas classes: os de partição das equações e os do tipo Newtoniano. Os métodos da primeira classe separam as equações e as variáveis em laços e as resolvem em cada laço até a convergência.

Entre os métodos de partição de equações está o método de ponto de bolha (Wang e Henke, 1966), que só é útil para sistemas cujos componentes tenham uma faixa relativamente estreita de razões de equilíbrio. Para os casos cuja faixa de volatilidades seja larga, Burningham e Otto (1967) propuseram que no já conhecido método de soma das taxas fosse empregado o algoritmo de Thomas para matrizes tridiagonais, na resolução das equações de balanço de massa modificadas.

Nos procedimentos baseados no método de Newton todas as variáveis são corrigidas simultaneamente a cada iteração. A principal diferença entre este tipo de métodos é o modo como as equações são organizadas. Naphtali e Sandholm (1971) propuseram que as equações de balanço de massa e energia e as razões de equilíbrio fossem agrupadas por estágios e então linearizadas. Deste modo, a matriz jacobiana do sistema de equações resultante tem uma estrutura de blocos tridiagonal, o que permite uma solução fácil das equações usando o algoritmo de Thomas. Um programa computacional que utiliza o método de Naphtali-Sandholm é ilustrado em Fredenslund *et alli* (1977). Já Ishii e Otto (1973) propuseram a resolução das equações para correções na temperatura, vazões de líquido e vapor e composições. Além disto, na linearização das equações somente as derivadas parciais de influência dominante são calculadas. Da mesma forma, Ferraris (1983) também propôs um novo arranjo para as equações linearizadas e variáveis a serem calculadas.

Os métodos Newtonianos têm sido usados em muitos simuladores comerciais (Taylor e Lucia, 1995). Os métodos de partição de equações podem apresentar sérios problemas de convergência de uma aplicação para outra, o que restringe o seu emprego.

2.2.1 – Método de Naphtali – Sandholm

Neste método as equações de balanço de massa por componente e de energia e de equilíbrio são agrupadas por estágio e linearizadas. A matriz jacobiana obtida tem uma estrutura de bloco tridiagonal, o que permite o emprego de um algoritmo simples para a resolução. O fluxograma de aplicação do método de Naphtali-Sandholm é apresentado no Apêndice A.

Um procedimento iterativo é empregado para reduzir as discrepâncias das equações M, E e H a valores muito próximos de zero. O procedimento iterativo é descrito a seguir.

Sejam os vetores:

$$\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_j, \dots, \mathbf{X}_{nst}]^T,$$

e

$$\mathbf{F} = [\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_j, \dots, \mathbf{F}_{nst}]^T,$$

\mathbf{X}_j é o vetor das variáveis no prato j e \mathbf{F}_j é o vetor das funções no prato j .

$$\mathbf{X}_j = [v_{1,j}, v_{2,j}, \dots, v_{i,j}, \dots, v_{nc,j}, T_j, l_{1,j}, l_{2,j}, \dots, l_{i,j}, \dots, l_{nc,j}]^T$$

$$\mathbf{F}_j = [H_j, M_{1,j}, M_{2,j}, \dots, M_{i,j}, \dots, M_{nc,j}, E_{1,j}, E_{2,j}, \dots, E_{i,j}, \dots, E_{nc,j}]^T$$

Então as equações do modelo da coluna apresentadas atrás devem ser escritas na forma genérica:

$$\bar{\mathbf{F}}(\bar{\mathbf{x}}) = 0.$$

A matriz jacobiana é formada pelas derivadas parciais das funções com respeito as variáveis. Como apenas variáveis dos pratos, anterior e posterior têm influência, esta matriz tem uma estrutura de blocos tridiagonais. Sendo,

$$\bar{\mathbf{A}}_j = \left(\frac{\partial \bar{F}_j}{\partial X_{j-1}} \right),$$

$$\bar{\mathbf{B}}_j = \left(\frac{\partial \bar{F}_j}{\partial X_j} \right),$$

$$\bar{\mathbf{C}}_j = \left(\frac{\partial \bar{F}_j}{\partial X_{j+1}} \right),$$

ter-se-á:

$$\left(\frac{\partial \overline{F}}{\partial X}\right) = \begin{bmatrix} \overline{B}_1 & \overline{C}_1 & 0 & \dots & \dots & 0 \\ \overline{A}_2 & \overline{B}_2 & \overline{C}_2 & 0 & \dots & \dots & \cdot \\ 0 & \overline{A}_3 & \overline{B}_3 & \overline{C}_3 & 0 & \dots & \dots & \cdot \\ \cdot & 0 & \ddots & \ddots & \ddots & 0 & \dots & \dots & \cdot \\ \cdot & \dots & 0 & \ddots & \ddots & \ddots & 0 & \dots & \dots & \cdot \\ \cdot & \dots & \dots & 0 & \ddots & \ddots & \ddots & 0 & \dots & \cdot \\ \cdot & \dots & \dots & \dots & 0 & \ddots & \ddots & \ddots & 0 & \cdot \\ \cdot & \dots & \dots & \dots & \dots & 0 & \overline{A}_{n-1} & \overline{B}_{n-1} & \overline{C}_{n-1} & \cdot \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & \overline{A}_n & \overline{B}_n & \cdot \end{bmatrix} \quad (2.19)$$

A cada iteração calcula-se o valor da correção a ser feita nas variáveis que estão sendo determinadas resolvendo o sistema de equações:

$$\Delta X^{(k)} = - \left[\left(\frac{\partial \overline{F}}{\partial X} \right)^{-1} \right]^{(k)} F^{(k)} \quad (2.20)$$

O valor da variável corrigida é calculado então por meio da seguinte equação:

$$X^{(k+1)} = X^{(k)} + \Delta X^{(k)} \quad (2.21)$$

Sabe-se que a matriz jacobiana tem uma estrutura de bloco tridiagonal, isto facilita o cálculo de ΔX , uma vez que se pode então empregar o algoritmo de Thomas para sua obtenção.

2.2.2 – Cálculo das Propriedades Termodinâmicas

Embora não figure explicitamente no modelo apresentado, o cálculo das propriedades termodinâmicas é indispensável para a simulação de colunas de destilação. Para ilustrar esta importância, dois casos serão considerados na simulação da coluna acima descrita.

No primeiro caso as propriedades termodinâmicas serão calculadas considerando as fases líquida e vapor como misturas ideais.

Quando as considerações de solução ideal e gás ideal são feitas o cálculo das propriedades termodinâmicas é drasticamente facilitado. Neste caso, as constantes de equilíbrio podem ser calculadas pela equação de Raoult, sendo a pressão de saturação obtida pela equação de Antoine (Yaws e Yang, 1989). As entalpias de líquido e de vapor podem ser calculadas por relações empíricas cuja única dependência funcional é em relação à temperatura. Yaws (1991) apresenta os parâmetros empíricos para o cálculo da entalpia de líquido por meio da equação:

$$\Delta H^L = \left(AT + \frac{BT^2}{2} + \frac{CT^3}{3} \right)_{T_1}^{T_2}$$

Já Yaws *et alli* (1988) apresentam os parâmetros para o cálculo da entalpia de vapor através da equação:

$$\Delta H^V = \left(aT + \frac{bT^2}{2} + \frac{cT^3}{3} + \frac{dT^4}{4} \right)_{T_1}^{T_2}$$

Como a utilização do método de Newton exige o cálculo das derivadas parciais relativas as variáveis que estão sendo obtidas. A hipótese de misturas ideais implica na redução do número de derivadas que devem ser calculadas, uma vez que apenas as derivadas parciais relativas à temperatura necessitam ser calculadas, desde que a pressão seja constante. Além disto, as derivadas parciais podem ser calculadas analiticamente com facilidade.

No segundo caso as propriedades termodinâmicas do sistema propano-propeno serão calculadas usando um pacote termodinâmico baseado na equação de Peng-Robinson (Zemp, 1995). Os passos para a aplicação da equação de Peng-Robinson são mostrados no Apêndice B.

Neste caso, além da temperatura e da pressão, as constantes de equilíbrio também sofrem a influência da composição. Do mesmo modo, as entalpias também não dependem apenas da temperatura, como no caso anterior, mas dependem ainda das composições e da

pressão. Assim sendo, um número maior de derivadas parciais necessita ser calculado. Além disto, as derivadas parciais relativas as composições são complicadas de ser obtidas analiticamente, sendo necessária a aplicação de métodos numéricos para seu cálculo.

2.2.3 – Especificações da Coluna

O método de Naphtali-Sandholm foi empregado para a simulação do modelo de uma coluna de destilação com condensador total (Figura 2.5) ou com condensador parcial (Figura 2.4).

A Tabela 2.1 mostra as especificações das colunas, tal como sugerida por Seferlis e Hrymak (1994).

Tabela 2.1 – Especificações da coluna.

Número de pratos	175
Prato de alimentação	61
Composição da alimentação	Propano: $x_2 = 0,1027$ Propeno: $x_1 = 0,8973$
Vazão de alimentação (Mmol / dia)	1,0734
Temperatura da alimentação (°C)	46,11

As variáveis fixadas, de acordo com a análise de graus de liberdade foram: a pressão, a razão de refluxo e a vazão de destilado. Estes valores são apresentados na Tabela 2.2.

Tabela 2.2 – Variáveis fixadas.

Pressão de operação (kPa)	1.860,60
Razão de refluxo	19,7
Vazão de destilado (Mmol / dia)	0,965

Em todas as simulações feitas as estimativas iniciais foram obtidas assumindo perfis lineares na coluna. O critério de convergência estabelecido foi:

$$\sum_{j=1}^{nsl} \left\{ (H_j)^2 + \sum_{i=1}^{nc} \left[(M_{i,j})^2 + (E_{i,j})^2 \right] \right\} \leq 5 \times 10^{-4}$$

2.2.4 – Resultados da Simulação

a. Misturas Ideais

O resultados da simulação para a coluna com condensador total são ilustrados nas Figuras 2.6 e 2.7. A Figura 2.6 ilustra os perfis de composição e a Figura 2.7 o perfil de temperatura na coluna.

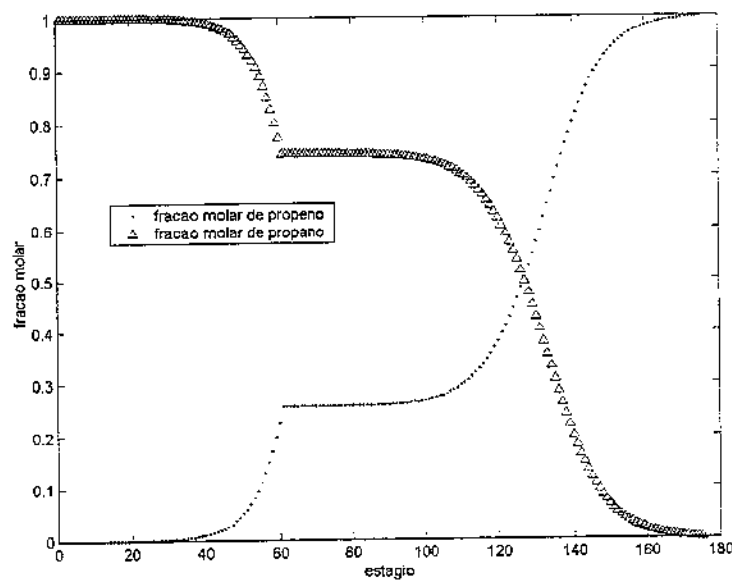


Figura 2.6 – Perfis de composição obtidos para condensador total (misturas ideais).

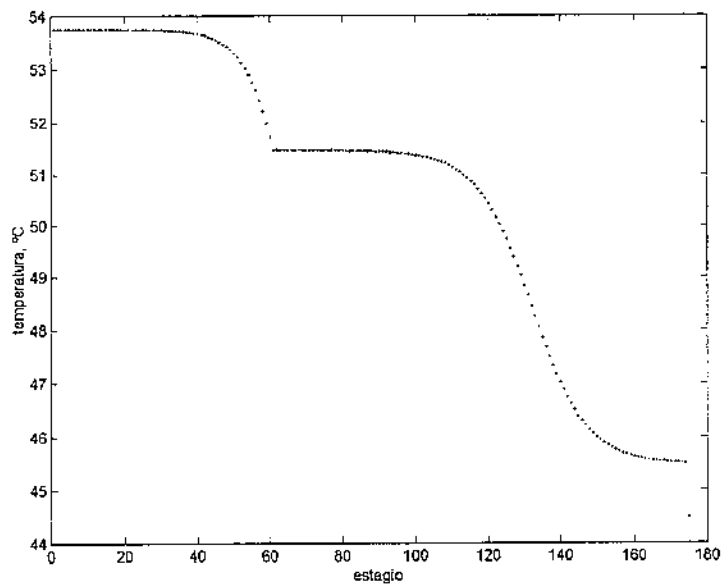


Figura 2.7 – Perfil de temperatura obtido para condensador total (misturas ideais).

Os resultados obtidos na simulação da coluna operando com condensador parcial são ilustrados nas Figuras 2.8 (perfis de composição) e 2.9 (perfil de temperatura).

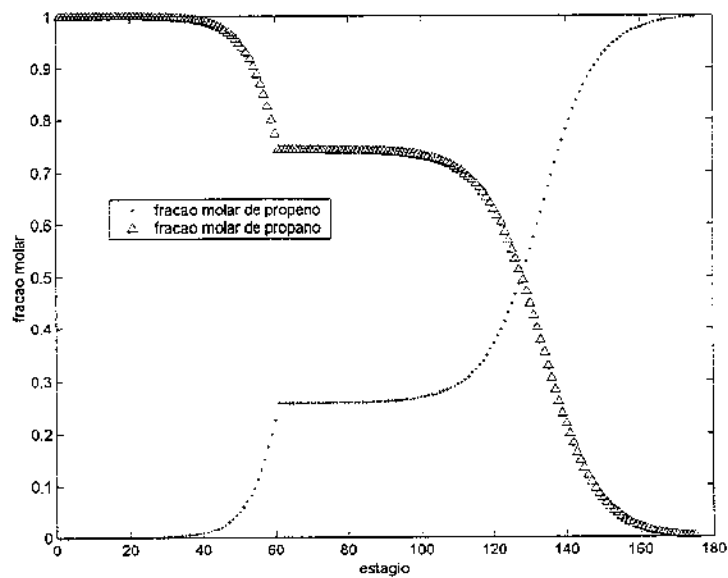


Figura 2.8 – Perfis de composição obtidos para condensador parcial (misturas ideais).

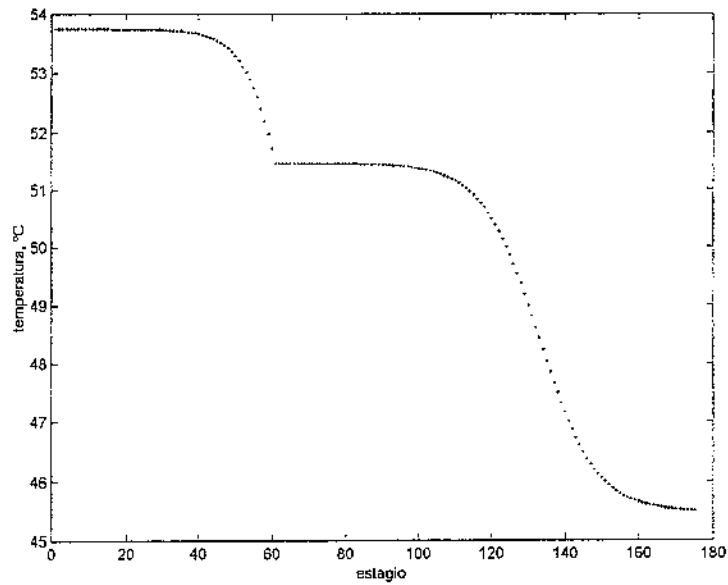


Figura 2.9 – Perfil de temperatura obtido para condensador parcial (misturas ideais).

b. Equação de Peng-Robinson

Os perfis de composição obtidos na simulação da coluna operando com condensador total são ilustrados na Figura 2.10 e o respectivo perfil de temperatura é apresentado na Figura 2.11.

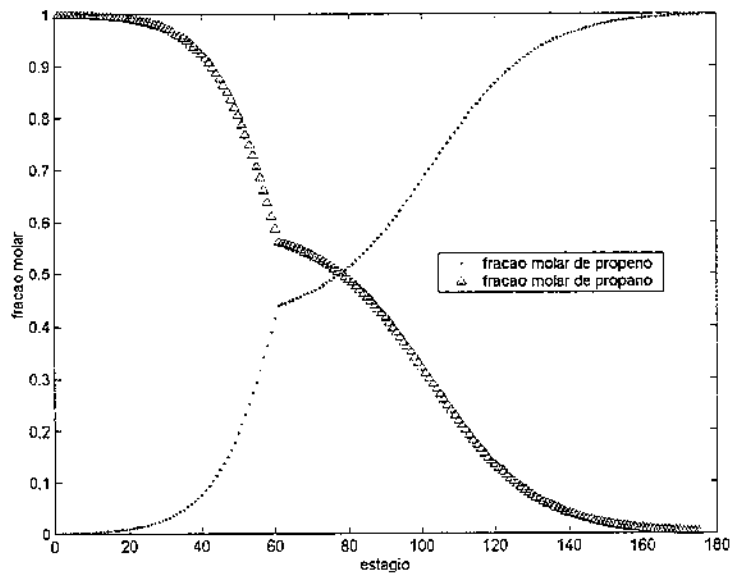


Figura 2.10 – Perfis de composição obtidos para condensador total (Peng-Robinson).

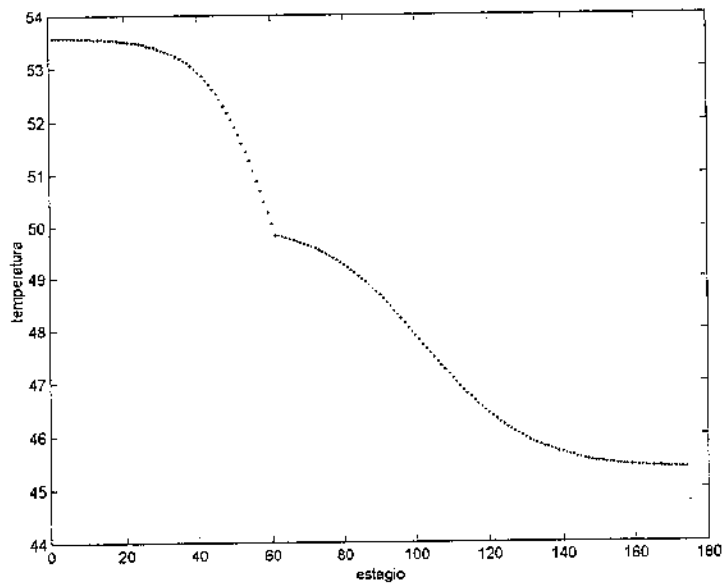


Figura 2.11 – Perfil de temperatura obtido para condensador total (Peng-Robinson).

A Figura 2.12 ilustra os perfis de composição obtidos uma vez que o condensador é do tipo parcial, enquanto que a Figura 2.13 apresenta o perfil de temperatura obtido neste caso.

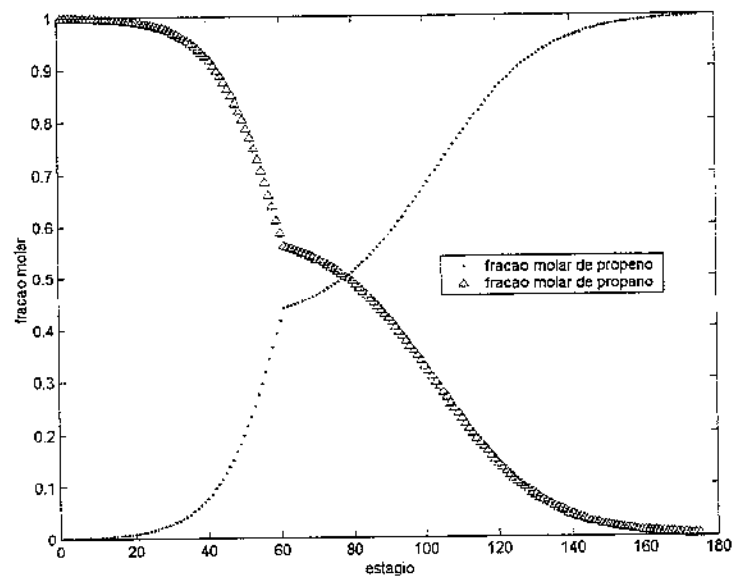


Figura 2.12 – Perfis de composição obtidos para condensador parcial (Peng-Robinson).

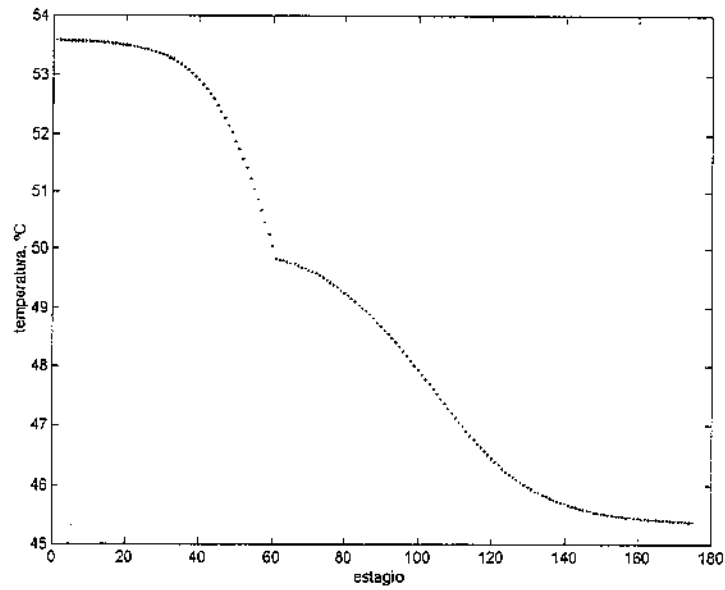


Figura 2.13 – Perfil de temperatura obtido para condensador parcial (Peng-Robinson).

Na Tabela 2.3 são apresentados os tempos computacionais e os números de iterações requeridos para a simulação da coluna operando com condensador total.

Tabela 2.3 – Tempos computacionais e número de iterações para coluna com condensador total.

Modelo usado nos cálculos termodinâmicos	Tempo de CPU (s)*	Número de iterações
Misturas ideais	0,60	10
Peng-Robinson	3,35	14

* SUN SPARC station 20

Na Tabela 2.4 são mostrados os tempos computacionais e os números de iterações requeridos na simulação da coluna com condensador parcial.

Tabela 2.4 – Tempos computacionais e número de iterações para coluna com condensador parcial.

Modelo usado nos cálculos termodinâmicos	Tempo de CPU (s)*	Número de iterações
Misturas ideais	1,12	21
Peng-Robinson	3,58	16

* SUN SPARC station 20

Na Tabela 2.5 são comparados os resultados obtidos pelos dois modelos termodinâmicos no caso de condensador parcial. Os valores obtidos pela equação de Peng-Robinson são considerados como referência. Para esta comparação foram calculados os somatórios dos erros quadráticos de cada fase, usando a seguinte equação:

$$E^f = \sum_{j=1}^{nst} \left[\left(z_{1,j,calc}^f - z_{1,j,ref}^f \right)^2 + \left(z_{2,j,calc}^f - z_{2,j,ref}^f \right)^2 \right] \quad (2.22)$$

onde $z_{i,j,calc}^f$ é a fração molar do componente i no prato j na fase f calculada, $z_{i,j,ref}^f$ é a fração molar do componente i no prato j na fase f referencial e E^f é o somatório do erro quadrático da fase f . O somatório do erro quadrático da temperatura é definido por:

$$E^T = \sum_{j=1}^{nst} \left[\left(T_{j,calc} - T_{j,ref} \right)^2 \right] \quad (2.23)$$

onde $T_{j,calc}$ é a temperatura calculada para o prato j e $T_{j,exato}$ é a temperatura de referência para o prato j .

Tabela 2.5 – Comparação dos resultados obtidos para coluna com condensador parcial.

Modelo termodinâmico usado	E^l	E^v	E^T
Peng-Robinson	-	-	-
Misturas ideais	21,4106	19,5330	772,975

Na Tabela 2.6 são comparados os resultados obtidos pelos dois modelos termodinâmicos no caso da coluna com condensador total.

Tabela 2.6 – Comparação dos resultados obtidos para coluna com condensador total.

Modelo termodinâmico usado	E^l	E^v	E^T
Peng-Robinson	-	-	-
Misturas ideais	21,3800	19,5046	771,861

2.3 – Conclusões

A importância da escolha do modelo termodinâmico pode ser verificada comparando as Figuras 2.6 e 2.10, 2.7 e 2.11, 2.8 e 2.12 e 2.9 e 2.13. Pelas Tabelas 2.5 e 2.6 pode-se perceber que os maiores desvios são referentes à temperatura. Os resultados também mostram uma diferença entre os perfis de temperatura da coluna operando com condensador total e parcial. A diferença é devida à temperatura do condensador. Esta diferença é previsível uma vez que um condensador total deve operar numa temperatura inferior à temperatura de ebulição do componente menos volátil. Enquanto o condensador parcial deve operar em uma temperatura algo acima do ponto de ebulição deste componente.

3 – Emprego de Redes Neurais Artificiais na Predição de Propriedades Termodinâmicas

3.1 – Introdução

Embora tenha uma história de cerca de cinquenta anos, apenas nos últimos quinze anos a utilização de redes neurais artificiais tem-se destacado. Isto se deve aos novos tipos de redes neurais, que possibilitaram sua utilização nas mais diversas áreas e aos novos algoritmos de treinamento que reduzem o tempo utilizado nesta etapa. Por estes motivos, as redes neurais estão sendo empregadas com sucesso nas mais diversas áreas. No âmbito da engenharia química a aplicação de redes neurais tem-se destacado em quatro grandes áreas: modelagem empírica, controle de processos, predição e detecção de falhas (Fontes, 2001).

Este capítulo introduz o emprego de redes neurais artificiais visando seu uso na predição de propriedades termodinâmicas no equilíbrio líquido-vapor (VLE).

A bibliografia referente ao emprego de redes neurais artificiais na predição de propriedades termodinâmicas é muito restrita. Sharma *et alli* (1999) usaram redes neurais artificiais em predições de equilíbrio líquido-vapor de sistemas binários. Estes usavam como entradas da rede neural artificial a pressão e a temperatura e obtinham como resultados a fração molar do componente em cada fase. Os autores concluíram que o emprego de redes neurais artificiais pode levar a resultados tão bons quanto ou até melhores que os preditos por equações de estado, mesmo empregando um conjunto limitado de dados. Iliuta *et alli* (2000) empregaram redes neurais artificiais na predição de equilíbrio líquido-vapor de sistemas com dois solventes e um eletrólito. Foram empregadas duas redes neurais: a primeira predizia a fração molar na fase vapor e a temperatura de equilíbrio e a segunda a fração molar na fase vapor e a pressão total.

As redes neurais artificiais surgiram da pesquisa sobre inteligência artificial cujo objetivo era modelar o comportamento do cérebro humano. Deste modo, uma entrada numérica é apresentada à rede de neurônios. Estes interagem matematicamente entre si e determinam uma resposta que corresponde a um padrão de entrada-saída pré-estabelecido.

Os neurônios das redes neurais são organizados em camadas. Todas as redes neurais apresentam uma camada de entrada, uma de saída e pelo menos uma camada escondida (Quantrille e Liu, 1991). As conexões entre os neurônios da rede neural artificial têm importância fundamental, para a retenção da informação adquirida. Outra fonte de retenção de informação da rede neural artificial é através da magnitude do sinal que percorre a rede neural. Quantrille e Liu (1991) citam que uma vez que a rede neural armazena conhecimento através das várias conexões entre neurônios, os efeitos causados por dados incoerentes ou com ruídos são distribuídos paralelamente pela rede e a influência da resposta de um neurônio não se torna muito significativa.

3.2 – Fundamentos de Redes Neurais Artificiais

3.2.1 – Modelo de um Neurônio

Os neurônios que compõe uma rede neural artificial apresentam uma estrutura conforme a ilustrada na Figura 3.1.

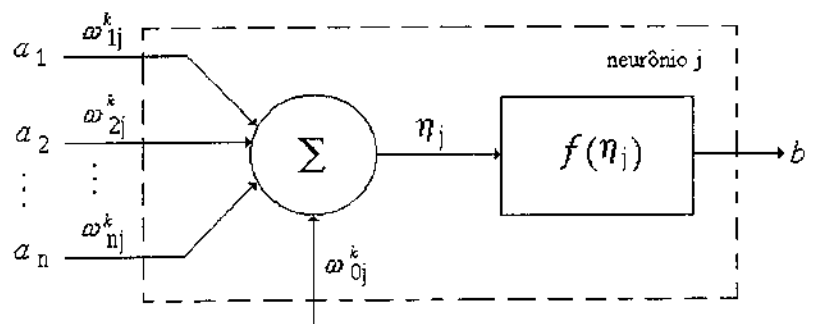


Figura 3.1 – Modelo do j-ésimo neurônio da camada k.

O primeiro elemento, é o vetor de entrada $\mathbf{a} = [a_1, a_2, \dots, a_n]^T$. Em função do vetor de entrada \mathbf{a} o neurônio j gera uma resposta, b_j . Os componentes de \mathbf{a} podem ser as respostas de outros neurônios, assim como a resposta b_j poderá fazer parte do vetor de entrada de outro neurônio.

O vetor $\boldsymbol{\omega}_j^k = [\omega_{1j}^k, \omega_{2j}^k, \dots, \omega_{nj}^k]^T$, compreende os pesos sinápticos do neurônio j da camada k , que são ajustados durante o treinamento da rede neural. O escalar ω_{0j}^k , chamado de limiar de ativação, é usado para controlar a ativação do neurônio. Estes elementos são manipulados de acordo com a equação:

$$\eta_j = \boldsymbol{\omega}_j^k \cdot \mathbf{a} + \omega_{0j}^k \quad (3.1)$$

O escalar η_j , por sua vez é alimentado à função de ativação, f , levando à saída b_j . A forma funcional de $f(\eta_j)$ pode ser escolhida de uma larga gama de funções matemáticas e pode variar nas diversas camadas da rede neural. Normalmente a forma funcional das camadas de entrada e de saída é linear, enquanto que a das camadas intermediárias comumente é uma função sigmóide, tal como

$$f(x) = \frac{1}{1 + e^{-x}}, \quad (3.2)$$

ou

$$f(x) = \frac{2}{1 + e^{-2x}} - 1. \quad (3.3)$$

Estas formas funcionais proporcionam treinamento mais eficiente, porque elas são monótonas. Além disso, elas têm a vantagem de serem contínuas e limitadas, entre 0 e 1, no caso da equação 3.2 e entre -1 e +1, no caso da equação 3.3. Elas são, ainda, diferenciáveis condição essencial para uma função de ativação (Demuth e Beale, 1998).

Usando a função sigmóide descrita pela equação 3.2, por exemplo, a saída b_j será dada por:

$$b_j = \frac{1}{1 + e^{-\eta_j}} \quad (3.4)$$

A equação 3.4 representa o último bloco ilustrado na Figura 3.1. As funções de ativação do tipo sigmóide são responsáveis pela inclusão de não linearidade ao modelo da rede neural. Este é um motivo pela qual este tipo de equações é muito empregado em modelagem empírica.

3.2.2 – Arquitetura de Redes Neurais Artificiais

Uma rede neural é composta por vários neurônios operando em paralelo e a resposta da rede é determinada pelas conexões entre neurônios. A maneira pela qual estes neurônios serão conectados é designada pela arquitetura da rede. Na maioria das aplicações, a rede neural é composta por mais de uma camada de neurônios. Geralmente redes neurais com maior número de camadas são capazes de representar com maior precisão o padrão de entrada-saída desejado.

O estudo da arquitetura mais conveniente para uma rede neural envolve a determinação do número de camadas e da função de ativação de cada uma delas, do número de neurônios em cada camada e como cada neurônio será ligado aos demais.

O emprego de redes neurais multicamadas possibilita a escolha entre três tipos de conexões: em conexões intracamadas a saída de um neurônio de uma camada é introduzida como entrada em outro da mesma camada; as conexões intercamadas em que a saída de um neurônio é alimentada aos neurônios de outras camadas; e as conexões recorrentes, na qual a saída de um neurônio é novamente introduzida a este. Estes tipos de conexões são ilustrados na Figura 3.2.

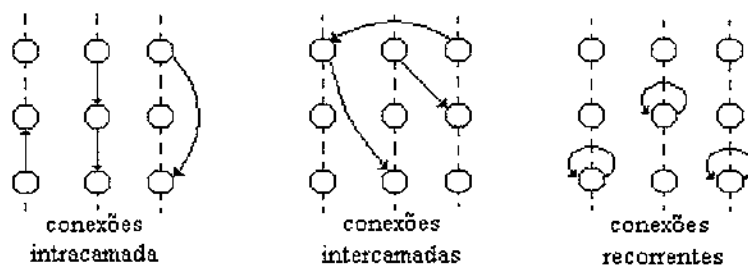


Figura 3.2 – Tipos de conexões.

Em engenharia, as conexões intercadas têm vantagens sobre os demais tipos (Quantrille e Liu, 1991). O emprego de conexões intercadas possibilita a escolha entre dois esquemas de conexão: as diretas ou as retrógradas. Nas conexões diretas, as entradas são introduzidas na primeira camada, estas por sua vez transmitem o sinal em direção a última camada, que emite a resposta. Enquanto nas conexões retrógradas, o sinal é transmitido no sentido inverso, isto é, parte da camada de saída em direção à camada de entrada. Ambos tipos de conexão estão ilustrados na Figura 3.3.

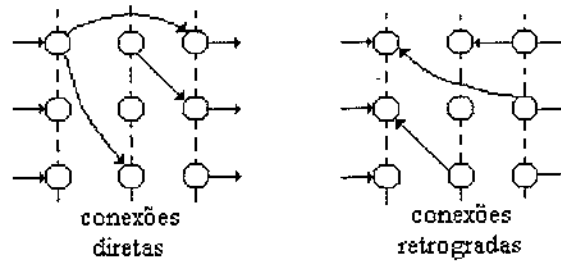


Figura 3.3 – Tipos de conexões intercadas.

Segundo Hunt *et alli* (1992) uma rede direta com três camadas, cuja não-linearidade seja descrita pela equação 3.1 e que tenha estrutura completamente conectada, pode representar qualquer função contínua. Uma rede deste tipo é ilustrada na Figura 3.4.

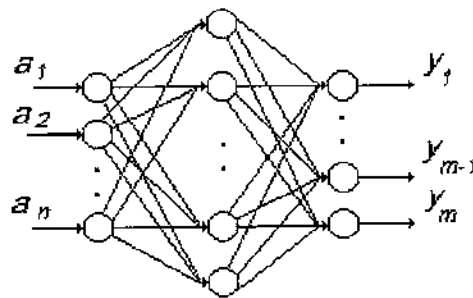


Figura 3.4 – Rede direta completamente conectada.

Estruturas como a da Figura 3.4 são muito úteis por serem simples e fáceis de implementar (Quantrille e Liu, 1991). Frequentemente a função de ativação da última camada

de uma rede neural como a ilustrada acima é uma função linear. Desta forma, o cálculo do vetor saída, \mathbf{y} , de uma rede neural como aquela é feito empregando-se as equações a seguir:

$$u_j = f(\boldsymbol{\omega}_j^h \cdot \mathbf{a} + \omega_{0j}^h) \text{ e} \quad (3.5)$$

$$\mathbf{y} = \boldsymbol{\omega}_j^o \cdot \mathbf{u} + \omega_{0o}^o, \quad (3.6)$$

onde $\boldsymbol{\omega}_j^h$ é o vetor dos pesos sinápticos referentes à camada intermediária do neurônio j , $\boldsymbol{\omega}_j^o$ é o vetor dos pesos sinápticos referentes a camada de saída do neurônio j , ω_{0j}^h é o limiar de ativação do neurônio j da camada intermediária e ω_{0o}^o é o limiar de ativação do neurônio o da camada de saída. Sendo $\mathbf{u} = [u_1, u_2, \dots, u_n]^T$.

Uma vez escolhida a estrutura da rede deve-se determinar o número de neurônios necessários para que a rede neural seja capaz de representar bem o padrão estabelecido durante o treinamento. Para uma rede neural semelhante à ilustrada na Figura 3.4, o número de neurônios na camada de entrada deve ser igual ao número de variáveis de entrada e o número de neurônios na camada de saída, igual ao número de variáveis de saída. Já a determinação do número de neurônios na camada intermediária é mais complexa. Um número inadequado de neurônios na rede pode levar a resultados inconvenientes.

Várias técnicas surgiram tentando resolver o problema da escolha do número ideal de neurônios na camada intermediária de uma rede neural. Fujita (1998) propôs um método estatístico para a estimação do número de neurônios. Já Zorzetto *et alli* (1999) propõe uma técnica de tentativa e erro, que se aproveita do conhecimento da tendência de redução do erro da rede.

Outras técnicas ainda, não buscam o número ideal de neurônios na camada intermediária de uma rede neural. Nestes métodos usa-se um grande número de neurônios, porém a boa generalização é garantida por uma técnica, chamada de regularização, que restringe os pesos sinápticos a valores reduzidos (Mackay, 1992).

3.2.3 – Utilização de Redes Neurais Artificiais

Antes da utilização de uma rede neural artificial para representar um determinado padrão de entrada-saída, os seguintes passos devem ser seguidos:

- 1) obtenção de um conjunto de padrões de entrada-saída;
- 2) divisão do conjunto em dois subconjuntos, um para o treinamento, um para a validação e um para a fase de testes;
- 3) determinação da estrutura da rede neural;
- 4) treinamento da rede neural. Nesta fase os pesos sinápticos e os limiares de ativação são ajustados até que um sinal de entrada produza uma saída desejada. O subconjunto de treinamento é empregado nesta fase. Um algoritmo de treinamento adequado deve ser escolhido;
- 5) generalização. Esta fase ocorre juntamente com a fase de treinamento, o subconjunto de validação é empregado para verificar se a rede neural não está memorizando os padrões usados para o treinamento. A memorização é um processo característico de redes com neurônios em excesso, por isso a determinação do número de neurônios adequado. A técnica de regularização é empregada nesta fase;
- 6) fase de testes. O último dos subconjuntos é empregado para comparar resultados obtidos empregando outros algoritmos de treinamento. Isto é, os passos 4 e 5 são repetidos para outros algoritmos de treinamento. O algoritmo que melhor representar os padrões do subconjunto de testes deve ser escolhido.

Então, a rede neural obtida representará o comportamento de entrada – saída conforme a precisão requerida. A rede já pode ser empregada para os fins desejados.

3.2.4 – Vantagens e Limitações das Redes Neurais Artificiais

As redes neurais não devem ser tomadas como uma verdade absoluta, aplicável a qualquer situação. Embora muito úteis, em alguns casos, sua utilização em outros casos tem sérias restrições. O conjunto de dados disponível deve ser suficientemente extenso para garantir a convergência e a estabilidade. Além disso, estes dados devem estar igualmente

distribuídos sobre toda a faixa de valores que se deseja prever. Outro problema é relacionado ao grande dispêndio de tempo para o treinamento de uma rede neural. Alguns algoritmos de treinamento que tentam acelerar o treinamento acabam requerendo maiores quantidades de memória computacional.

O crescente interesse pelo emprego de redes neurais é vinculado às diversas vantagens que elas proporcionam. A capacidade de adaptação é uma destas vantagens. Caso a rede esteja fora de seu ambiente de validade ela pode ser treinada para adaptar-se a este novo ambiente. Além disso, uma rede neural treinada pode simular uma entrada muito rapidamente. Isto proporciona à rede a possibilidade de ser empregada em técnicas on-line. A capacidade de representar relações não-lineares é outra propriedade muito apreciável das redes neurais.

Outra vantagem característica de redes neurais é a possibilidade de se calcular as derivadas das saídas em relação às entradas. Esta característica é muito apreciável, pois muitas relações funcionais são muito complicadas de serem representadas analiticamente. O cálculo da derivada é feito como se segue.

Sendo:

$$Df(x) \equiv \frac{df(x)}{dx},$$

então a derivada da saída y_o em relação à entrada a_i é obtida por:

$$\frac{\partial y_o}{\partial a_i} = \sum_{j=1}^{n_i} \omega_{oj}^o [Df(o_j) \omega_{ij}^h], \quad (3.7)$$

onde n_i é o número de neurônios na camada intermediária.

3.3 – Técnicas de Treinamento

O processo de treinamento, também chamado aprendizado, de uma rede neural artificial, consiste no ajuste dos pesos sinápticos e dos limiares de ativação de maneira que um padrão de entrada-saída seja obedecido.

Existem dois tipos de treinamento: o supervisionado; no qual um “*supervisor*” externo monitora o treinamento, e o não-supervisionado; no qual a rede neural é treinada sem nenhum controle externo. O treinamento supervisionado é o mais usado e existem vários algoritmos aplicando este tipo de treinamento. Dentre estes, o algoritmo de treinamento por correção de erros é o mais utilizado. Neste os pesos sinápticos e os limiares de ativação são ajustados proporcionalmente ao vetor dos erros das saídas, \mathbf{e} , dado por:

$$\mathbf{e}_i = \mathbf{o}_i - \mathbf{t}_i, \quad (3.8)$$

onde \mathbf{o} é o vetor dos valores calculados e \mathbf{t} é o vetor dos valores desejados.

O algoritmo de treinamento por correção de erros é útil apenas em redes neurais com uma única camada. Este problema acabou retardando a difusão desta ferramenta computacional. Porém com o surgimento da técnica de retropropagação redes neurais com mais de uma camada puderam ser empregadas. A única exigência para o emprego desta técnica é que a rede neural a ser treinada possua apenas ligações intercamadas diretas.

O algoritmo de retropropagação sofreu várias modificações e melhorias, porém os passos básicos são (Quantrille e Liu, 1991):

- 1) Apresenta-se uma entrada específica e mede-se a saída atual;
- 2) Comparam-se as saídas; atual e desejada e calcula-se o erro quantitativo entre estas saídas; a equação usada para calcular este erro é chamada de função de desempenho e pode assumir várias formas;
- 3) Minimiza-se este erro ajustando os pesos sinápticos e os limiares de ativação.
 - a) Ajustam-se os pesos sinápticos e os limiares de ativação a partir da última em direção primeira camada, o ajuste é feito proporcionalmente ao erro em cada camada.

O algoritmo de retropropagação mais simples é o de gradiente descendente. Neste procedimento, os pesos sinápticos e os limiares de ativação são ajustados na direção em que o erro diminui mais rapidamente. O treinamento de uma rede neural mediante este algoritmo leva muito tempo. Isto ocasionou a obtenção de procedimentos que empregam técnicas heurísticas e de otimização numérica para acelerar o treinamento. A técnica de Levenberg-

Marquardt é empregada com este propósito. O algoritmo obtido é muito eficiente e merece maior aprofundamento.

3.3.1 – Algoritmo de Levenberg-Marquardt

O método de Newton é uma alternativa para acelerar o treinamento de redes neurais. Uma vez que se sabe que este alcança convergência rapidamente, o processo de atualização dos pesos sinápticos e dos limiares de ativação de redes neurais pode aproveitar esta característica. Aplicando o método de Newton ao treinamento de redes neurais, obtém-se:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{H}_k^{-1} \mathbf{g}_k, \quad (3.9)$$

aqui, \mathbf{x}_{k+1} é o vetor que contém os pesos sinápticos e os limiares de ativação corrigidos, \mathbf{x}_k contém os valores atuais destes parâmetros, \mathbf{H}_k^{-1} é a matriz Hessiana da função desempenho e \mathbf{g}_k é o gradiente atual.

O cálculo exato da matriz Hessiana é muito complexo e requer um esforço computacional elevado. No algoritmo de Levenberg-Marquardt a matriz Hessiana não é calculada, mas sim uma aproximação que é adequada a funções desempenho que tenham forma quadrática, descrita matematicamente como:

$$\mathbf{H} \approx \mathbf{J}^T \mathbf{J},$$

onde \mathbf{J} é a matriz Jacobiano da função performance. Já o gradiente, \mathbf{g}_k , é calculado como:

$$\mathbf{g}_k = \mathbf{J}^T \mathbf{e}.$$

A matriz Jacobiano é calculada com muito mais facilidade do que a matriz Hessiana. À matriz Hessiana aproximada acrescenta-se um termo para garantir que a função desempenho será sempre reduzida após cada iteração do algoritmo. Desta forma:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}] \mathbf{J}^T \mathbf{e}, \quad (3.10)$$

onde: \mathbf{I} é a matriz Identidade e μ é uma variável que garante a redução da função performance.

Embora este algoritmo seja rápido, ele requer uma memória computacional elevada. Este problema pode ser amenizado dividindo a matriz Jacobiano em duas submatrizes iguais. As submatrizes são manipuladas como descrito abaixo:

$$\mathbf{H} = \mathbf{J}^T \mathbf{J} = \begin{pmatrix} \mathbf{J}_1^T \\ \mathbf{J}_2^T \end{pmatrix} (\mathbf{J}_1, \mathbf{J}_2) = \mathbf{J}_1^T \mathbf{J}_1 + \mathbf{J}_2^T \mathbf{J}_2$$

Após cada subtermo da matriz ter sido computado a parte correspondente da matriz Jacobiano pode ser eliminada liberando memória.

Demuth e Beale (1998) compararam o tempo gasto no treinamento de uma rede neural específica usando vários algoritmos; o algoritmo de Levenberg-Marquardt foi o mais veloz.

3.4 – Melhorando a Generalização

O objetivo do treinamento de uma rede neural é fazer não só que está presente 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 do treinamento. A memorização é característica de redes superdimensionadas. O emprego de redes neurais 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 deste número exato de neurônios é difícil. Por este motivo surgiram outros métodos para melhorar a generalização de redes neurais. O método de regularização está entre estes e será descrito a seguir.

3.4.1 – Regularização

Regularização é o nome dado ao método que, para melhorar a generalização de uma rede neural, restringe os valores dos pesos sinápticos e dos limiares de ativação a números reduzidos. Esta restrição leva a rede a responder suavemente, reduzindo os proble-

mas causados pela memorização. Deste modo até mesmo uma rede superdimensionada será capaz de obedecer bem a um determinado padrão de entrada-saída.

O emprego de regularização requer o emprego de uma função desempenho específica. Na maioria das vezes a função desempenho empregada, tem a seguinte forma:

$$F = E_D = \sum_{i=1}^n (\mathbf{t}_i - \mathbf{o}_i)^2, \quad (3.11)$$

onde n é o número de padrões de entrada-saída usados no treinamento. Porém no caso de emprego de regularização usa-se a seguinte função desempenho:

$$F = \beta E_D + \alpha E_W, \quad (3.12)$$

onde E_W é o somatório dos quadrados dos pesos sinápticos da rede e α e β são chamados de parâmetros de regularização.

O ajuste dos valores de α e β é o principal problema na implementação deste método. Para contornar este problema a lei de Bayes foi empregada, dando origem à técnica ora chamada de regularização Bayesiana.

3.4.2 – Otimização dos Parâmetros de Regularização pela Lei de Bayes – Regularização Bayesiana

A lei de Bayes é usada para avaliar a possibilidade de ocorrência de um dado evento,

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}, \quad (3.13)$$

onde $P(A|B)$ representa a probabilidade do evento A ocorrer se o evento B já aconteceu, $P(B|A)$ é a probabilidade do evento B ocorrer dado que o evento A já aconteceu, $P(A)$ é a probabilidade do evento A ocorrer e $P(B)$ é a probabilidade do evento B ocorrer.

Considerando os pesos sinápticos e limiares de ativação da rede como variáveis randômicas e conhecendo os padrões de entrada-saída que serão usados no treinamento, a função densidade para os pesos sinápticos e limiares de ativação pode ser obtida pela lei de Bayes:

$$P(\mathbf{F} | D, \alpha, \beta, M) = \frac{P(D | \mathbf{F}, \beta, M)P(\mathbf{F} | \alpha, M)}{P(D | \alpha, \beta, M)}, \quad (3.14)$$

onde D representa o conjunto de padrões de treinamento, M é o modelo sugerido para a rede neural e \mathbf{F} é a matriz que contém os todos os pesos sinápticos e os limiares de ativação.

Assumindo que os desvios nos dados de treinamento e a distribuição inicial dos pesos sinápticos e limiares de ativação são Gaussianos, as probabilidades podem ser dadas por:

$$P(D | \mathbf{F}, \beta, M) = \frac{1}{Z_D(\beta)} \exp(-\beta E_D) \quad \text{e} \quad (3.15)$$

$$P(\mathbf{F} | \alpha, M) = \frac{1}{Z_w(\alpha)} \exp(-\alpha E_w) \quad (3.16)$$

onde $Z_D(\beta) = (\pi/\beta)^{n/2}$, $Z_w = (\pi/\alpha)^{N/2}$, N é o número total de parâmetros e n é o número total de padrões usados no treinamento. Substituindo na equação 3.14, tem-se:

$$P(\mathbf{F} | D, \alpha, \beta, M) = \frac{1}{Z_F(\alpha, \beta)} \exp(-F(w)). \quad (3.17)$$

Assim, quando os pesos ótimos forem obtidos $P(\mathbf{F}|D, \alpha, \beta, M)$ será máximo e a função desempenho, F , será mínima.

Resolvendo a equação 3.14 para $P(D|\alpha, \beta, M)$, obtém-se:

$$P(D | \alpha, \beta, M) = \frac{Z_F(\alpha, \beta)}{Z_D(\beta)Z_w(\alpha)}. \quad (3.18)$$

Na equação 3.18 somente $Z_F(\alpha, \beta)$ é desconhecido. Porém, ele pode ser estimado usando expansão em séries de Taylor. Como a função desempenho tem a forma quadrática em uma pequena área próxima do ponto mínimo F_w pode ser expandido ao redor do ponto mínimo de \mathbf{F}^{MP} , onde o gradiente é zero. Obtendo, desta maneira:

$$Z_F \approx (2\pi)^{N/2} \det(\mathbf{H}^{MP})^{-1/2} \exp(-F(\mathbf{F}^{MP})),$$

onde \mathbf{H} é a matriz Hessiana da função desempenho. Substituindo na equação 3.18 os valores ótimos de α e β no ponto mínimo podem ser obtidos; para isto deve-se derivar os logaritmos e igualá-los a zero, obtendo-se:

$$\alpha^{MP} = \frac{\gamma}{2E_w(\mathbf{F}^{MP})} \text{ e } \beta^{MP} = \frac{N - \gamma}{2E_D(\mathbf{F}^{MP})},$$

onde $\gamma = N - 2\alpha^{MP} \text{tr}(\mathbf{H}^{MP})^{-1}$ é chamado de número efetivo de parâmetros. Ele representa o número de parâmetros efetivamente utilizados na redução da função desempenho. Este valor pode variar de zero a N . Uma diferença muito grande entre γ e N é sinal alguns neurônios da camada intermediária podem ser eliminados.

A técnica de regularização Bayesiana, assim como a técnica de Levenberg-Marquardt para o treinamento necessita do cálculo da matriz Jacobiano da função desempenho. Desta forma, a técnica de regularização Bayesiana pode ser implementada juntamente com o algoritmo de Levenberg-Marquardt sem aumentar muito o esforço computacional. Os melhores resultados são obtidos se o conjunto de padrões de entrada-saída usado no treinamento for normalizado na faixa de $[-1,1]$ (ou outra região similar) (Foresse e Hagan, 1997).

Foresse e Hagan (1997) descrevem os passos necessários para a otimização Bayesiana dos parâmetros da função desempenho usando a técnica de Levenberg-Marquardt:

- 1) Inicializar α , β e os pesos sinápticos e limiares de ativação;

- 2) Computar um passo do algoritmo de Levenberg – Marquardt para minimizar o valor da função desempenho;
- 3) Calcular o número efetivo de parâmetros, usando a seguinte aproximação para a matriz Hessiana, $\mathbf{H} \approx 2\beta\mathbf{J}^T\mathbf{J} + 2\alpha\mathbf{I}_N$;
- 4) Calcular α e β ;
- 5) Calcular os passos 2 até 4 até a convergência.

Cada vez que α e β são calculados, o valor da função desempenho é alterado e o ponto mínimo é modificado. Desta forma, a cada nova estimativa os parâmetros da função desempenho serão mais precisos. Isto ocorrerá até o momento em que a precisão for suficientemente elevada, de maneira que a função desempenho permanecerá inalterada. Neste momento, a convergência é obtida e o treinamento da rede neural irá cessar.

3.5 – Redes Neurais Artificiais na Predição de Propriedades Termodinâmicas

A predição de dados termodinâmicos por meio de redes neurais foi testada para uma mistura propano-propeno. As propriedades que serão preditas pelas redes neurais são necessárias na simulação de colunas de destilação.

Para isto foram empregadas três redes neurais: a primeira, **R1**, é responsável pela obtenção das constantes de equilíbrio líquido-vapor a partir das frações molares nas fases líquida e vapor do componente mais volátil e da temperatura. Embora pela regra de fases de Gibbs o estado de equilíbrio líquido-vapor de um sistema binário seja definido por apenas duas variáveis intensivas, no procedimento iterativo do método de simulação da coluna de destilação as composições nas fases líquida e vapor e a temperatura são corrigidas a cada iteração. Por este motivo, admite-se que exista um estado de equilíbrio líquido-vapor hipotético nestas condições. Por meio da segunda, **R2**, deseja-se calcular a entalpia da fase líquida tendo como entradas a fração molar do componente mais volátil na fase líquida e a temperatura. A última rede, **R3**, é utilizada na predição da entalpia de vapor a partir da fração molar do componente mais volátil na fase vapor e da temperatura. As derivadas parciais necessárias no método iterativo devem ser calculadas por estas redes neurais. Estas redes são ilustradas na Figura 3.5.

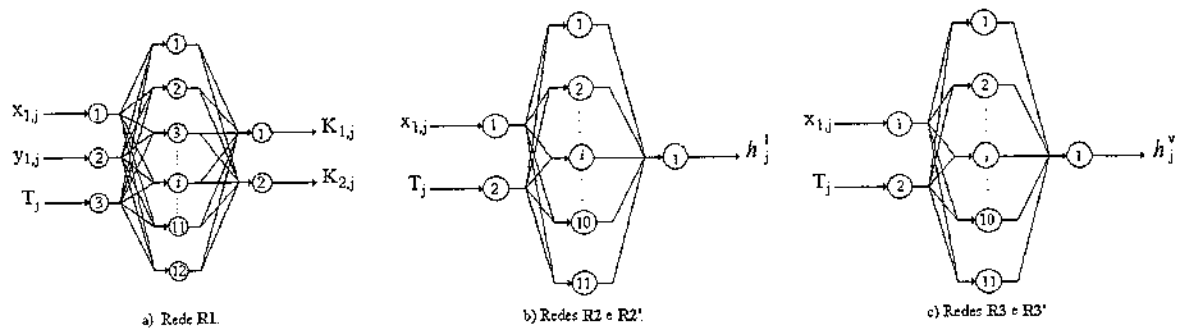


Figura 3.5 – Redes neurais empregadas.

Para o treinamento das redes neurais foi empregado o software MATLAB. Este possui um pacote que utiliza o algoritmo de Levenberg-Marquardt juntamente com a técnica de Regularização Bayesiana para o treinamento de redes neurais. Este pacote foi empregado no treinamento das redes neurais acima descritas.

Os dados usados no treinamento das redes neurais foram gerados a partir da equação de Peng-Robinson (Zemp, 1995). Estes dados cobrem toda a faixa de composições, uma faixa de temperatura compreendida entre 45,367 °C (ponto de ebulição do propeno) e 53,593 °C (ponto de ebulição do propano) e são válidos para a pressão de 1.860,60 kPa. Um conjunto de quatrocentos padrões de entrada-saída foi gerado. Deste total, trezentos padrões foram usados no treinamento e os cem restantes foram empregados para testar as redes neurais obtidas. Para o caso de coluna operando com condensador total as redes **R2** e **R3**, foram substituídas por duas redes neurais alternativas, **R2'** e **R3'**, para as quais o limite inferior da temperatura foi reduzido para 43,938 °C (temperatura abaixo do ponto de ebulição do componente mais volátil) e o número de padrões subiu para quatrocentos e quatro, sendo que o conjunto de treinamento aumentou para 303 padrões de entrada-saída.

3.5.1 – Normalização dos Padrões de Entrada-Saída

Os padrões de entrada-saída usados no treinamento foram normalizados na faixa de [-1,1]. Todas as variáveis, tanto de entrada quanto de saída, foram normalizadas de acordo com a equação abaixo:

$$p_n = \frac{2(p - p_{\min})}{p_{\max} - p_{\min}} - 1 \quad (3.19)$$

onde p_n é o valor da variável normalizada, p é o valor original da variável, p_{min} é um valor mínimo de referência da variável e p_{max} é um valor máximo de referência da variável.

Desta forma quando a rede neural for usada para predizer um certo valor, as entradas devem ser normalizadas de acordo com a equação 3.19. Já as saídas devem ser revertidas aos valores corretos por meio da equação:

$$p = 0,5(p_n + 1)(p_{max} - p_{min}) + p_{min} \quad (3.20)$$

3.5.2 – Resultados Obtidos

Os valores obtidos após o treinamento, para o número efetivo de parâmetros e para a somatória dos quadrados dos pesos sinápticos, são apresentados na Tabela 3.1. Estes valores indicam que as redes estão bem dimensionadas.

Tabela 3.1 – Parâmetros de treinamento.

Rede	Número efetivo de parâmetros, γ	Somatória dos quadrados dos pesos sinápticos
R1	23,40	45,86
R2	9,210	36,87
R2'	19,30	130,1
R3	15,30	36,69
R3'	19,80	168,2

O desempenho das redes pode ser avaliado por meio dos erros obtidos no treinamento e os erros obtidos no teste das redes neurais. Os valores dos somatórios dos quadrados dos erros (SSE) encontram-se na Tabela 3.2.

Nas Figuras 3.6 e 3.7 são comparados os valores calculados pelas redes e aqueles obtidos pelo modelo de Peng-Robinson para o conjunto de padrões usados nos testes da rede R1.

Tabela 3.2 – Erros obtidos no treinamento e no teste das redes neurais.

Rede	Somatório dos quadrados dos erros do treinamento	Somatório dos quadrados dos erros do teste
R1	$1,1277 \times 10^{-8}$	$1,2931 \times 10^{-11}$
R2	$6,5521 \times 10^{-9}$	$4,1853 \times 10^{-4}$
R2'	$1,1030 \times 10^{-8}$	$7,3456 \times 10^{-4}$
R3	$6,7622 \times 10^{-10}$	$2,8096 \times 10^{-5}$
R3'	$4,4893 \times 10^{-9}$	$1,8644 \times 10^{-3}$

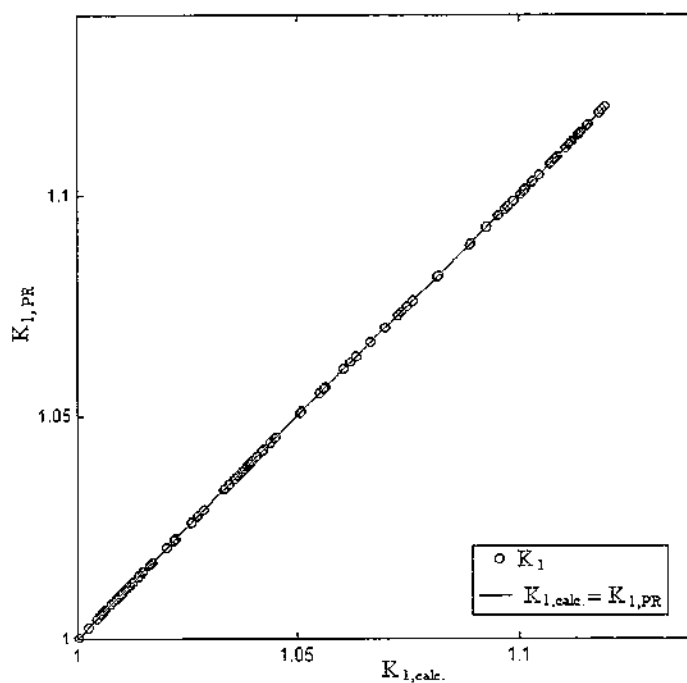


Figura 3.6 – Comparação entre os valores da constante de equilíbrio do propeno preditos por Peng-Robinson e os calculados pela rede neural R1.

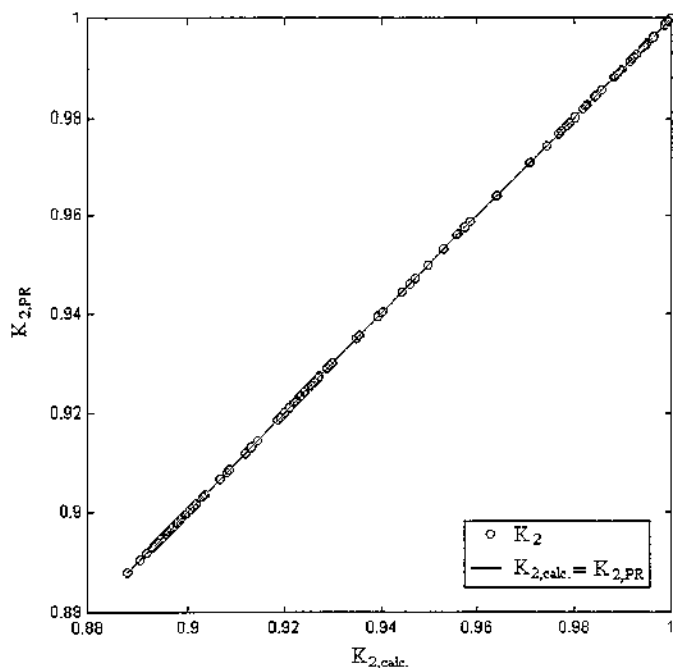


Figura 3.7 – Comparação entre os valores da constante de equilíbrio do propano preditos por Peng-Robinson e os calculados pela rede neural R1.

Os valores calculados pelas redes R2, R2', R3 e R3' são comparados com os calculados pela equação de Peng-Robinson, nas Figuras 3.8, 3.9, 3.10 e 3.11, respectivamente.

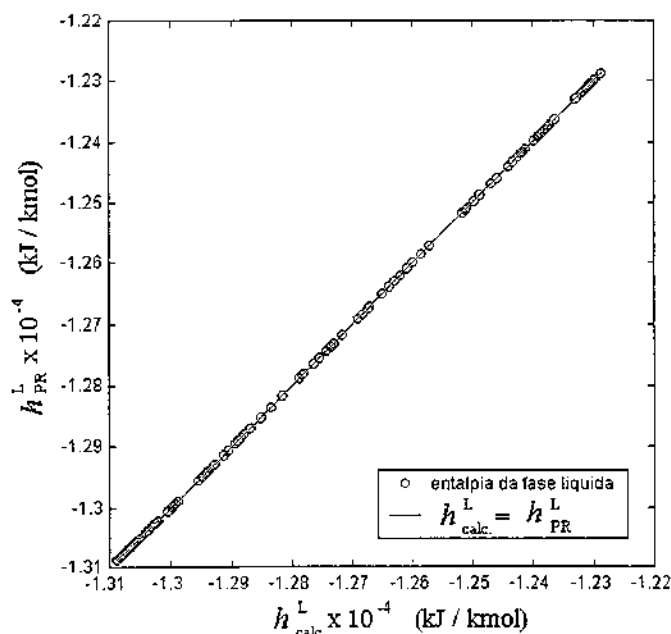


Figura 3.8 – Comparação entre os valores da entalpia da fase líquida preditos por Peng-Robinson e os calculados pela rede neural R2.

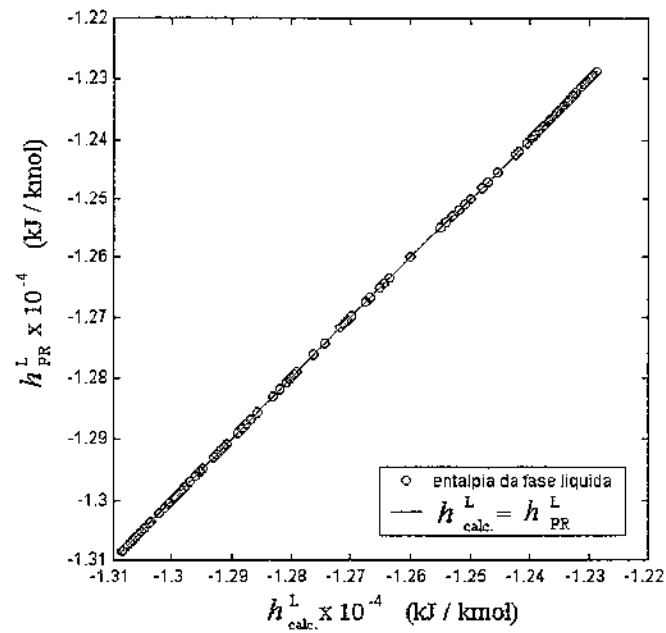


Figura 3.9 – Comparação entre os valores da entalpia da fase líquida preditos por Peng-Robinson e os calculados pela rede neural R2’.

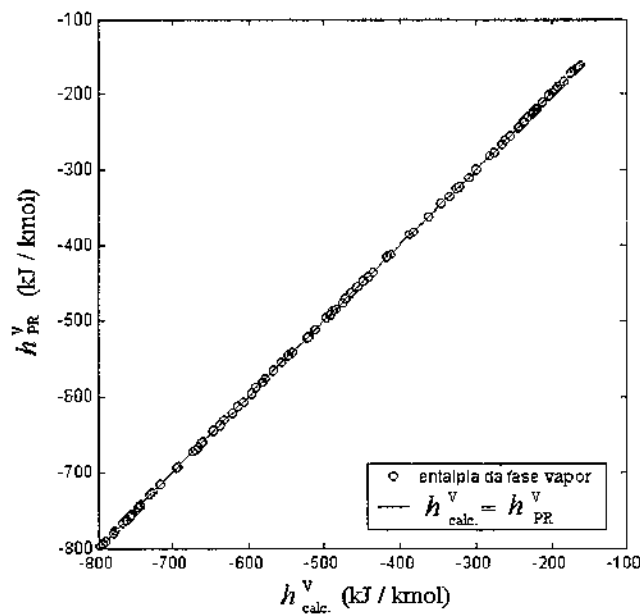


Figura 3.10 – Comparação entre os valores da entalpia da fase vapor preditos por Peng-Robinson e os calculados pela rede neural R3.

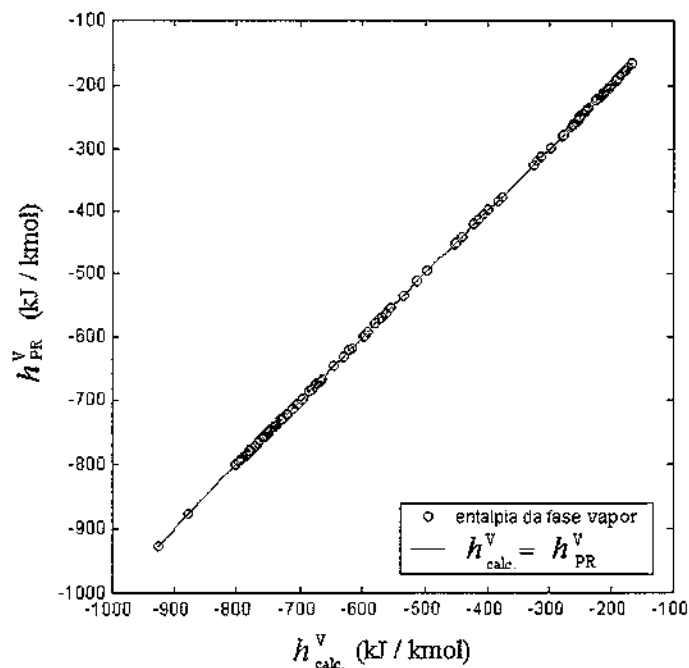


Figura 3.11 – Comparação entre os valores da entalpia da fase vapor preditos por Peng-Robinson e os calculados pela rede neural R3⁷.

Os parâmetros das redes são mostrados nas Tabelas 3.3, 3.4, 3.5, 3.6 e 3.7.

Tabela 3.3 – Parâmetros da rede R1.

neurônio	1	2	3	4	5	6	7	8	9	10	11	12
$\omega_1^{h^T}$	-0,2876	-1,2780	0,1701	0,2338	0,1731	-0,1375	0,2337	1,2690	0,2343	0,2342	0,2981	0,2344
$\omega_2^{h^T}$	0,8431	0,1839	-0,9188	-0,0391	-0,5045	-0,2102	-0,0391	0,9404	-0,0423	-0,0423	0,1085	-0,0422
$\omega_3^{h^T}$	0,9620	-0,3957	0,2960	-0,2424	1,5288	0,4702	-0,2425	0,2968	-0,2446	-0,2446	-0,7163	-0,2443
$\omega_{01}^{h^T}$	0,8516	1,5625	-2,3503	-0,0499	-1,3077	0,2069	-0,0498	1,4119	-0,0539	-0,0537	0,0426	-0,0538
ω_1^p	0,0668	1,7013	1,5234	-0,9170	1,0488	0,4008	0,9169	1,0115	-0,9152	-0,9152	-0,5462	-0,9154
ω_2^p	1,7226	0,8696	1,5069	-0,7259	-0,4032	0,900	0,7260	-0,4530	-0,7301	-0,7301	-0,5475	-0,7298
ω_{00}^p	-0,3462	-0,0807										

Tabela 3.4 - Parâmetros da rede R2.

neurônio	1	2	3	4	5	6	7	8	9	10	11
$\omega_1^{h^T}$	-0,2995	0,2203	-0,3123	0,2221	0,2177	0,2230	0,2200	0,2226	-0,3118	-0,6601	0,5013
$\omega_2^{h^T}$	-0,6000	0,5029	-0,6157	0,5051	0,4998	0,5061	0,5025	0,5057	-0,6151	0,1140	-0,4821
$\omega_{01}^{h^T}$	-0,2464	0,2670	-0,2405	0,2653	0,2693	0,2645	0,2672	0,2649	-0,2407	-1,1012	3,2297
ω_1^p	-1,3200	1,1735	-1,3307	1,1750	1,1719	1,1753	1,1733	1,1750	-1,3302	0,8292	-2,4971
ω_{00}^p	-0,0923										

Tabela 3.5 - Parâmetros da rede R2^o.

neurônio	1	2	3	4	5	6	7	8	9	10	11
ω_1^{hT}	3,9404	-0,5323	3,6707	2,7854	-1,3705	-0,5729	-0,6833	1,0524	-0,5354	0,2786	0,5910
ω_2^{hT}	2,7894	0,0041	0,2527	-0,1441	2,5872	-0,0249	-0,1052	-0,00234	1,8669	-0,1441	-0,6014
ω_{01}^{hT}	-2,8729	-0,2330	-2,7460	0,0485	2,4955	-0,2720	-0,3671	-0,3060	3,0180	0,0485	2,8766
ω_1^o	3,0919	1,1212	-2,2923	-0,3732	-2,2712	1,1657	1,3142	-0,7381	2,8181	-0,3732	-2,1274
ω_{0o}^oT	0,6230										

Tabela 3.6 - Parâmetros da rede R3.

neurônio	1	2	3	4	5	6	7	8	9	10	11
ω_1^{hT}	0,1710	0,0658	-1,4855	-0,5046	0,1955	1,3130	0,4099	1,0110	-0,2306	-0,3746	-0,4114
ω_2^{hT}	0,6245	0,5220	0,5818	-1,1634	0,6490	-1,1226	-1,5333	-0,8486	-0,6866	-0,8257	0,9097
ω_{01}^{hT}	0,0055	-0,0095	0,5326	-0,6511	0,0233	0,1094	-0,0883	0,7866	-0,0214	-0,0913	-2,8006
ω_1^o	0,8559	0,7074	0,5638	-1,6093	0,8864	-0,1678	0,9502	-0,4255	-1,0525	-1,2381	1,9714
ω_{0o}^oT	-0,1800										

Tabela 3.7 - Parâmetros da rede R3^o.

neurônio	1	2	3	4	5	6	7	8	9	10	11
ω_1^{hT}	-1,4238	-4,5747	0,2321	4,6512	-0,2863	1,7648	1,7133	-0,2709	0,0972	0,2321	0,2321
ω_2^{hT}	0,8451	-2,7350	-0,1509	0,7101	0,1395	-2,5615	0,7893	2,1430	1,3500	-0,1509	-0,1509
ω_{01}^{hT}	1,6590	3,5068	0,0602	-3,6089	-0,1599	-2,9760	-0,0710	2,9277	-2,8572	0,0602	0,0602
ω_1^o	0,7326	-3,0203	-0,4006	-2,4959	0,8779	2,4018	-1,3655	2,7450	2,9418	-0,4006	-0,4006
ω_{0o}^oT	0,4392										

3.6 – Conclusões

Como se pode verificar na Tabela 3.2, os somatórios dos quadrados dos erros das três redes são muito pequenos. Isto comprova a eficiência das redes neurais na predição destas propriedades termodinâmicas. As ordens de grandeza dos somatórios dos quadrados dos erros são proporcionais as ordens de grandeza dos valores usados no treinamento.

4 – Modelos Estáticos Híbridos de Colunas de Destilação

4.1 – Introdução

As técnicas de controle (controle preditivo com modelo) e de otimização on-line fazem uso intensivo de modelos (dinâmicos ou estáticos) do processo a controlar ou otimizar. Por esta razão têm sido feitos esforços para permitir obter modelos de processos cujo tempo computacional seja reduzido.

As técnicas mais usadas para a redução do tempo computacional associado à simulação de colunas de destilação de pratos (assim como de outros processos semelhantes) fazem uso ou de Modelos Reduzidos ou de Modelos Simplificados (Perregaard e Sørensen, 1992).

Os Modelos Reduzidos são basicamente modelos de ordem reduzida, ou seja, representados por um número substancialmente menor de equações em relação aos modelos “exatos” apresentados anteriormente. Já com os Modelos Simplificados procuram-se introduzir hipóteses que tornam mais simples, para a simulação, as equações do modelo. Neste sentido, o próprio Modelo de Estágio de Equilíbrio já é um Modelo Simplificado.

Uma das técnicas mais usadas para a redução da ordem de modelos de colunas de destilação de pratos é o método de Colocação Ortogonal Discreta (Simões, 2000).

Um dos aspectos freqüentemente considerados na simplificação de modelos de colunas de destilação por sua vez, diz respeito à redução do tempo computacional gasto nos cálculos termodinâmicos. Perregaard (1993) relata que o tempo necessário para o cálculo das propriedades termodinâmicas representa de 30 a 80% do tempo total da simulação. Já no caso da simulação da coluna de destilação apresentada na seção 2.2 a porcentagem de tempo necessária para estes cálculos foi de cerca de 64%. Este tempo foi calculado usando a ferramenta do compilador (Fortran Salford FTN95) que permite a obtenção do tempo computacional gasto em cada subrotina do programa computacional. Uma técnica de simplificação que visa a redução do tempo usado nos cálculos termodinâmicos é a utilização dos chamados “*modelos termodinâmicos locais*”. Os modelos termodinâmicos locais são conjuntos de equações simples usadas em vez dos modelos rigorosos, que necessitam de procedimentos iterativos para sua solução, dos modelos rigorosos, é substituído por expres-

sões simples. Estas equações são válidas para regiões limitadas de composição, temperatura e pressão. Quando as condições do sistema atingem os limites destas regiões é necessária uma atualização. Um dos problemas básicos encontrados na utilização de modelos locais é determinar o momento em que o modelo deve ser atualizado. Chimowitz *et alli* (1985) utilizam um procedimento de atualização automático. Perregaard (1993) propôs que a atualização fosse feita cada vez que a matriz jacobiana era calculada. Já Støren e Hertzberg (1994) usaram um modelo de erro local para diagnosticar a necessidade de atualização. Todos conseguiram acelerar os cálculos termodinâmicos em taxas consideráveis.

4.2 – Modelos Híbridos

A utilização de modelos híbridos é uma alternativa quando existem parâmetros de difícil obtenção através de técnicas de cálculo convencionais, como equações de estado. O emprego destas envolve cálculos iterativos que podem consumir muito tempo. Uma boa introdução sobre modelos híbridos é apresentada por Psychogios e Ungar (1992). Neste trabalho, os autores apresentam a utilização de um modelo híbrido para simular um bioreator. Para testar o modelo híbrido eles comparam os resultados com técnicas já conhecidas e com o emprego de um modelo “caixa preta” do bioreator. Quanto à comparação com o modelo “caixa preta” eles comentam que o modelo híbrido tem maior capacidade de interpolação e extrapolação, é de melhor compreensão e requer uma quantia significativa de padrões a menos para o treinamento.

4.3 – Metodologia Empregada

O modelo híbrido a ser usado neste trabalho é composto por duas partes: a primeira compreende as equações de balanços de massa e energia e as relações de equilíbrio, apresentadas na seção 2.1; a segunda é composta por três redes neurais responsáveis pelos cálculos termodinâmicos. A primeira rede neural, **R1**, terá como entradas a temperatura e as frações molares do componente mais leve nas fases líquida e vapor, e suas saídas serão as constantes de equilíbrio para ambos os componentes. A segunda rede, **R2**, terá como entradas a temperatura e a fração molar do componente mais leve na fase líquida, a saída da rede será a entalpia da corrente líquida. A terceira rede, **R3**, terá como entradas a temperatura e a fração molar do componente mais leve na fase vapor e sua saída será a entalpia da corrente

de vapor. Para o caso de coluna operando com condensador total as redes **R2** e **R3**, serão substituídos pelas redes **R2'** e **R3'**. Estas redes neurais foram preparadas anteriormente e seus resultados constam no Capítulo 3.

O sistema de equações linearizadas será resolvido pelo método de Naphtali-Sandholm, apresentado na seção 2.2.1. As derivadas parciais serão calculadas por meio de redes neurais. Sendo que apenas as derivadas parciais relativas à temperatura serão calculadas, uma vez que as demais são muito pequenas em relação a estas. Isto pode ser verificado na Tabela 4.1, onde são apresentadas as derivadas parciais calculadas durante a simulação da coluna de destilação com condensador parcial apresentada no Capítulo 2. O fluxograma de resolução do modelo híbrido da coluna de destilação é apresentado na Figura 4.1.

Tabela 4.1 – Comparação dos valores das derivadas parciais.

Derivada parcial em relação a variável no prato i	Prato		
	i=1	i=87	i=175
$\partial k_1 / \partial y_{i,1}$	$-3,4638 \times 10^{-4}$	$-7,2432 \times 10^{-6}$	$6,2939 \times 10^{-7}$
$\partial k_2 / \partial y_{i,1}$	$2,5213 \times 10^{-4}$	$8,2055 \times 10^{-6}$	$7,4050 \times 10^{-8}$
$\partial k_1 / \partial y_{i,2}$	$1,0061 \times 10^{-6}$	$1,0150 \times 10^{-5}$	$1,0687 \times 10^{-6}$
$\partial k_2 / \partial y_{i,2}$	$2,1126 \times 10^{-7}$	$-9,9521 \times 10^{-5}$	$-3,9249 \times 10^{-5}$
$\partial k_1 / \partial x_{i,1}$	$2,2002 \times 10^{-4}$	$3,3962 \times 10^{-5}$	$9,5289 \times 10^{-8}$
$\partial k_2 / \partial x_{i,1}$	$-5,1967 \times 10^{-6}$	$-4,4457 \times 10^{-5}$	$-3,7905 \times 10^{-6}$
$\partial k_1 / \partial x_{i,2}$	$-4,5740 \times 10^{-7}$	$-5,0858 \times 10^{-5}$	$-5,7056 \times 10^{-4}$
$\partial k_2 / \partial x_{i,2}$	$-2,0706 \times 10^{-8}$	$6,0767 \times 10^{-5}$	$1,4353 \times 10^{-3}$
$\partial k_1 / \partial T_i$	$1,4936 \times 10^{-2}$	$1,4899 \times 10^{-2}$	$1,4798 \times 10^{-2}$
$\partial k_2 / \partial T_i$	$1,3796 \times 10^{-2}$	$1,3661 \times 10^{-2}$	$1,3494 \times 10^{-2}$

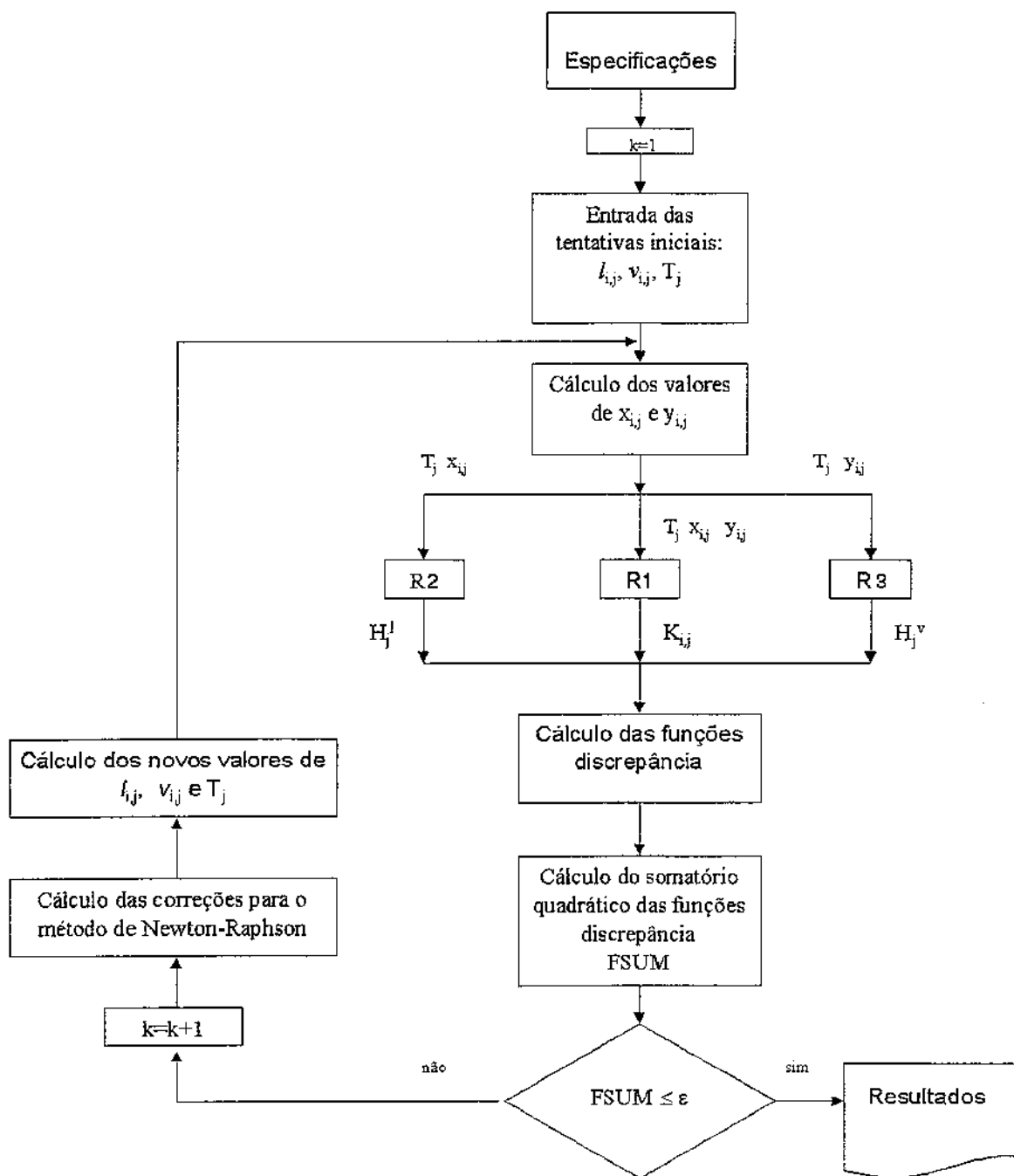


Figura 4.1 – Fluxograma de resolução do modelo híbrido.

4.4 – Resultados

O modelo híbrido foi testado para as duas situações, condensador total e condensador parcial. Inicialmente serão apresentados os resultados obtidos para a coluna com condensador parcial. As estimativas iniciais são as mesmas empregadas na seção 2.2.4.

4.4.1 – Condensador Parcial

Considerando as mesmas condições operacionais apresentadas na Figura 4.2 os perfis de composição obtidos pelo modelo híbrido são comparados com os resultados apresentados na seção 2.2.

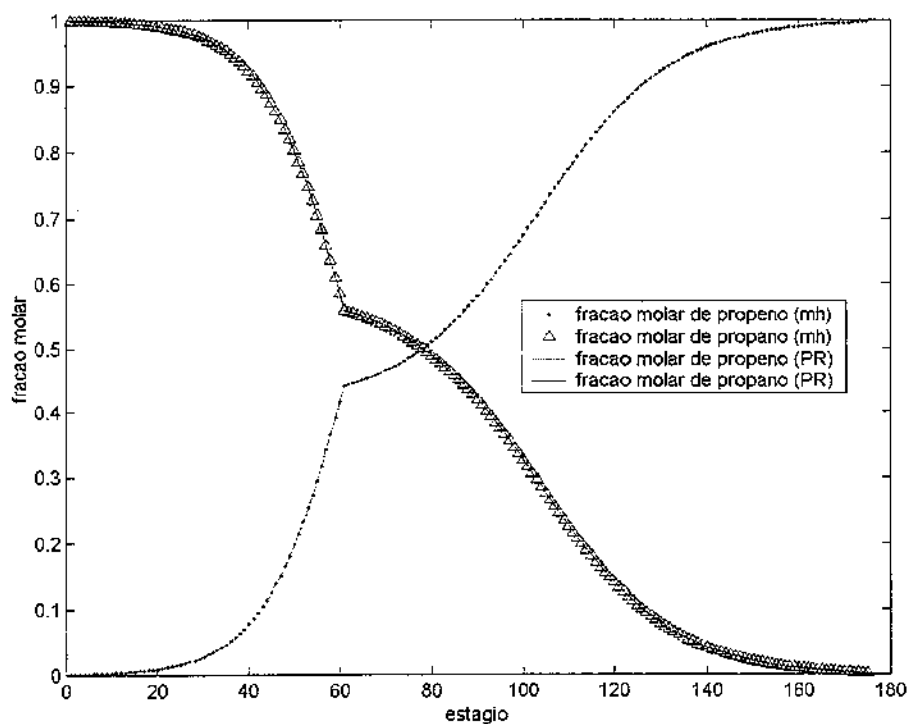


Figura 4.2 – Comparação dos perfis de composição obtidos para o modelo que utiliza as equações de Peng-Robinson (PR) e para o modelo híbrido (mh).

Já o perfil de temperatura obtido usando o modelo híbrido é comparado com os resultados de referência e apresentado na Figura 4.3.

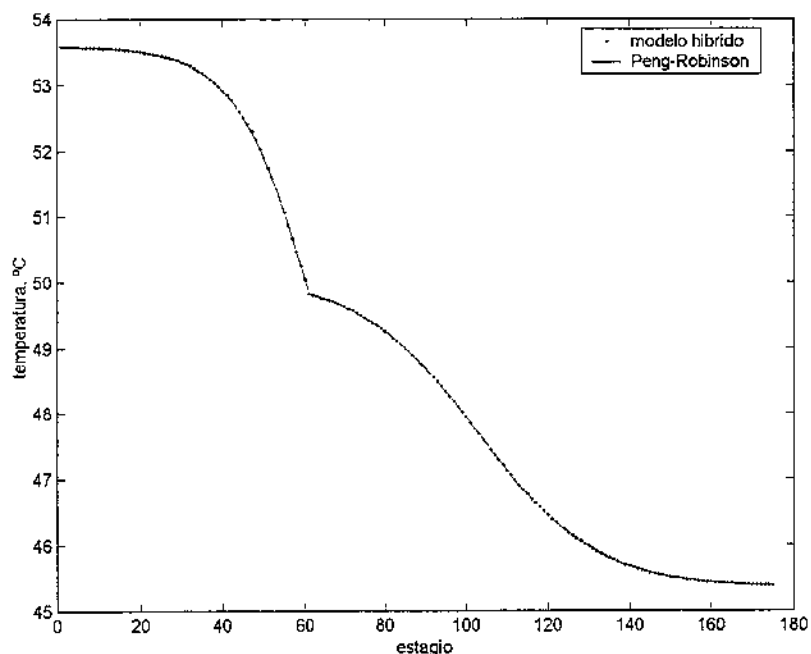


Figura 4.3 – Comparação dos perfis de temperatura obtidos para o modelo que utiliza as equações de Peng-Robinson e para o modelo híbrido.

Os tempos computacionais, os erros e o número de iterações resultantes da simulação são apresentados na Tabela 4.2. A solução do modelo que emprega as equações de Peng-Robinson foi usada como referência.

Tabela 4.2 – Resultados obtidos para coluna com condensador parcial.

Modelo usado nos cálculos termodinâmicos	Número de iterações	Somatório do Erro quadrático da fase líquida	Somatório do Erro quadrático da fase vapor	Somatório do Erro quadrático da temperatura	Tempo de CPU* (s)
Peng-Robinson	16	-	-	-	3,58
Modelo híbrido	12	$2,3602 \times 10^{-5}$	$2,1470 \times 10^{-5}$	$1,0457 \times 10^{-3}$	1,62

*SUN SPARC station 20

4.4.2 – Condensador Total

Os resultados obtidos pelo modelo híbrido para a coluna com condensador total são comparados aos que foram calculados na seção 2.2.4 e ilustrados nas Figuras 4.6 e 4.7. Na Tabela 4.3 são apresentados os resultados da simulação do modelo híbrido para a coluna com condensador total. Novamente os resultados obtidos usando a equação de Peng-Robinson são usados como referência.

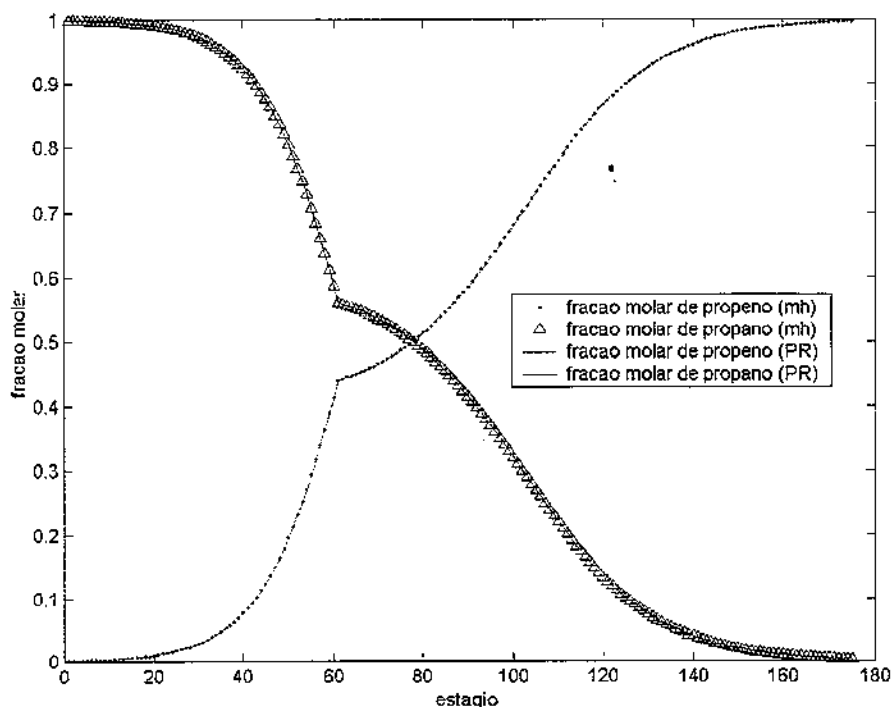


Figura 4.6 – Comparação dos perfis de composição obtidos para o modelo que utiliza as equações de Peng-Robinson (PR) e para o modelo híbrido (mh).

Tabela 4.3 – Resultados obtidos para coluna com condensador total.

Modelo usado nos cálculos termodinâmicos	Número de iterações	Somatório do Erro quadrático da fase líquida	Somatório do Erro quadrático da fase vapor	Somatório do Erro quadrático da temperatura	Tempo de CPU* (s)
Usando equações de Peng-Robinson	14	-	-	-	3,35
Modelo híbrido	10	$7,5546 \times 10^{-6}$	$7,3967 \times 10^{-6}$	$3,6317 \times 10^{-2}$	1,51

*SUN SPARC station 20

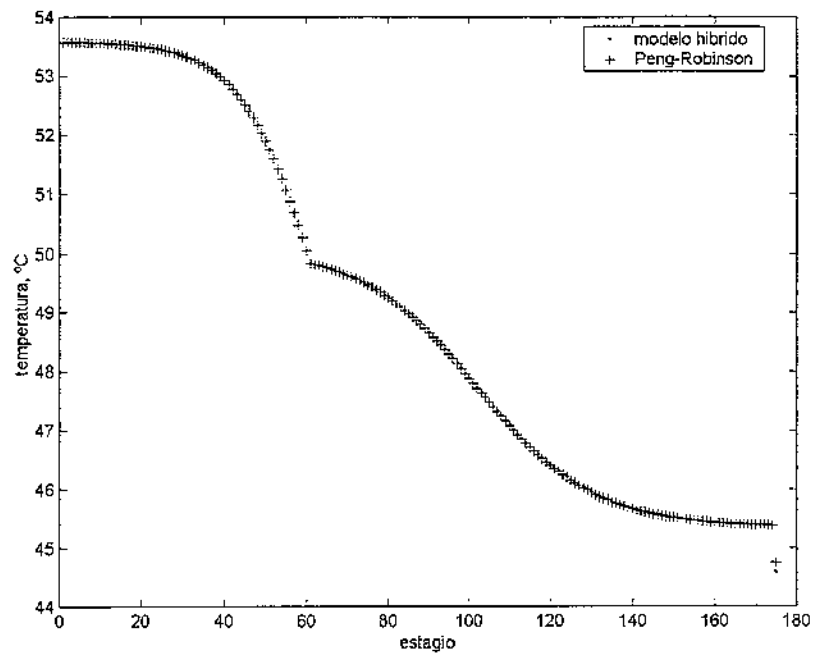


Figura 4.7 – Comparação dos perfis de temperatura obtidos para o modelo que utiliza as equações de Peng-Robinson (PR) e para o modelo híbrido (mh).

4.5 – Conclusões

Os resultados obtidos pelo modelo híbrido foram muito próximos dos valores obtidos pelo modelo original. Porém o tempo computacional requerido pelo modelo híbrido foi em média 55% menor.

5 – Discussão dos Resultados

O objetivo deste trabalho é testar o emprego de um modelo híbrido na simulação de colunas de destilação. Neste, as propriedades termodinâmicas deveriam ser calculadas por meio de redes neurais.

O emprego da técnica de Regularização Bayesiana no treinamento das redes neurais evitou o gasto de tempo no dimensionamento destas, para evitar o processo de memorização dos padrões usados no treinamento. As redes neurais obtidas tinham boas propriedades de generalização.

Já o treinamento das redes neurais é muito demorado. Porém com o surgimento de técnicas e computadores cada vez mais rápidos este tempo tende a ser reduzido.

Os resultados obtidos e apresentados no Capítulo quarto demonstram a viabilidade do emprego de redes neurais na predição de propriedades termodinâmicas para a simulação de colunas de destilação. Os resultados foram suficientemente precisos, tanto para simulação da coluna com condensador total quanto para condensador parcial. Além disto, o tempo computacional requerido na simulação foi reduzido significativamente.

5.1 – Sugestões para Trabalhos Futuros

1. Testar o emprego de redes neurais na simulação dinâmica de colunas de destilação para futura aplicação em técnicas de controle por modelo e otimização;
2. Usar modelos híbridos para a simulação (dinâmica e estática) de colunas de destilação com mais de dois componentes;
3. Empregar redes neurais para a predição das propriedades termodinâmicas usadas em modelos de ordem reduzida;

Referências Bibliográficas

- BURNINGHAN, D.W., OTTO, F.D. Which Computers Design for Absorbers? *Hydrocarbon Processing*, v.46, n.10, 1967.
- CHIMOWITZ, E.H., MACCHIETTO, S., ANDERSON, T.F. Dynamic Multicomponent Distillation Using Local Thermodynamic Models. *Chemical Engineering Science*, v.40, n.10, p. 1974-1978, 1985.
- DEMUTH, H., BEALE, M. Neural Network Toolbox User's Guide. Natick: The MathWorks, Inc., 2000.
- FERRARIS, G.B. A Powerful Improvement of the Global Newton-Raphson Method for Multistaged Multicomponent Separators. *Computers and Chemical Engineering*, v.7, n.2, p. 73-85, 1983.
- FONTES, C.H.O. *Análise e Controle de um Reator de Copolimerização*, Campinas: Faculdade de Engenharia Química, Universidade Estadual de Campinas, 2001. Tese (Doutorado)
- FORESSE, F.D., HAGAN, M.T. Gauss-Newton Aproximation to Bayesian Learning. In: *Proceedings of the 1997 International Joint Conference on Neural Networks*, p. 1930-1935, 1997.
- FREDENSLUND, A., GMEHLING, J., RASMUSSEN, P. Vapor-Liquid Equilibria Using UNIFAC, A Group Contribution Method. Amsterdam: Elsevier, 1977.

- FUJITA, O. Statistical Estimation of the Number of Hidden Units for Feedforward Neural Networks. *Neural Networks*, v. 11, p. 851-859, 1998.
- HUNT, K.J., SBARBARO, D., ZBIKOWSKI, R., GAWTHROP, P.J. Neural Networks for Control Systems – A Survey. *Automatica*, v.28, n.6, p. 1083-1112, 1992.
- ILIUTA, M.C., ILIUTA, I., LARACHI, F. Vapour-Liquid Equilibrium Data Analysis for Mixed Solvent-Electrolyte Systems Using Neural Network Models. *Chemical Engineering Science*, v.55, p. 2813-2825, 2000.
- ISHII, Y., OTTO, F.D. A General Algorithm for Multistage Multicomponent Separation Calculations. *The Canadian Journal of Chemical Engineering*, v.51, p. 601-606, 1973.
- MACKAY, D.J.C. Bayesian Interpolation. *Neural Computation*, v.4, n.3, p. 415-447, 1992.
- NAPHTALI, L.M., SANDHOLM, D.P. Multicomponent Separation Calculation by Linearization. *AIChE Journal*, v.17, n.1, p. 148-153, 1971.
- PERREGAARD, J., SØRENSEN, E.L., Simulation and Optimization of Chemical Processes: Numerical and Computacional Aspects. *Computers and Chemical Engineering*, Suppl. 16, 1992.
- PERREGAARD, J. Model Simplification and Reduction for Simulation and Optimization of Chemical Processes. *Computers and Chemical Engineering*, v.17, n.5/6, p. 465-483, 1993.

- PSICHOGIOS, D.C., UNGAR, L.H. A Hybrid Neural Network-First Principles Approach to Process Modeling. *AIChE Journal*, v.38, n.10, 1992.
- QUANTRILLE, T.E., LIU, Y.A. Artificial Intelligence in Chemical Engineering. San Diego: Academic Press, Inc., 1991.
- REID, R.C., PRAUZNITZ, J.M., POLING, B.E. Properties of Gases and Liquids. New York: McGraw-Hill, 1987.
- SEFERLIS, P., HRYMAK, A.N. Optimization of Distillation Units Using Collocation Models. *AIChE Journal*, v.40, n.5, p. 813-825, 1994.
- SHARMA, R., SINGHAL, D., GHOSH, R., DWIVEDI, A. Potential Applications of Artificial Neural Networks to Thermodynamics: Vapor-Liquid Equilibrium Predictions. *Computers and Chemical Engineering*, v.23, p. 385-390, 1999.
- STØREN, S., HERTZBERG, T., Local Thermodynamic Models Applied in Dynamic Process Simulation: A Simplified Approach. *Trans IChemE*, v. 72, Part A, 1994.
- TAYLOR, R., LUCIA, A. Modeling and Analysis of Multicomponent Separation Processes. In: *AIChE Symposium Series*, v.91, 1995.
- WANG, J.C., HENKE, G.E. Tridiagonal Matrix for Distillation. *Hydrocarbon Processing*, v.45, n.8, 1966.

YAWS, C.L., NI, H.M., CHIANG, P.Y. Heat Capacities for 700 Compounds. *Chemical Engineering*, v. 9, p. 91-98, 1988.

YAWS, C.L., YANG, H.C. To Estimate Vapor Pressure Easily. *Hydrocarbon Processing*, p. 65-68, 1989.

YAWS, C.L. Calculate Liquid Heat Capacity. *Hydrocarbon Processing*, p. 73-77, 1991.

ZEMP, R.J. Software de Propriedades Termodinâmicas. DESQ/FEQ, UNICAMP, 1995.

ZORZETTO, L.F.M., MACIEL FILHO, R., MACIEL, M.R.W. Process Modelling Development Through Artificial Neural Networks and Hybrid Models. *Computers and Chemical Engineering*, v. 24, p. 1355-1360, 2000.

Apêndice A – Método de Naphtali-Sandholm

Fluxograma para emprego do método de Naphtali-Sandholm.

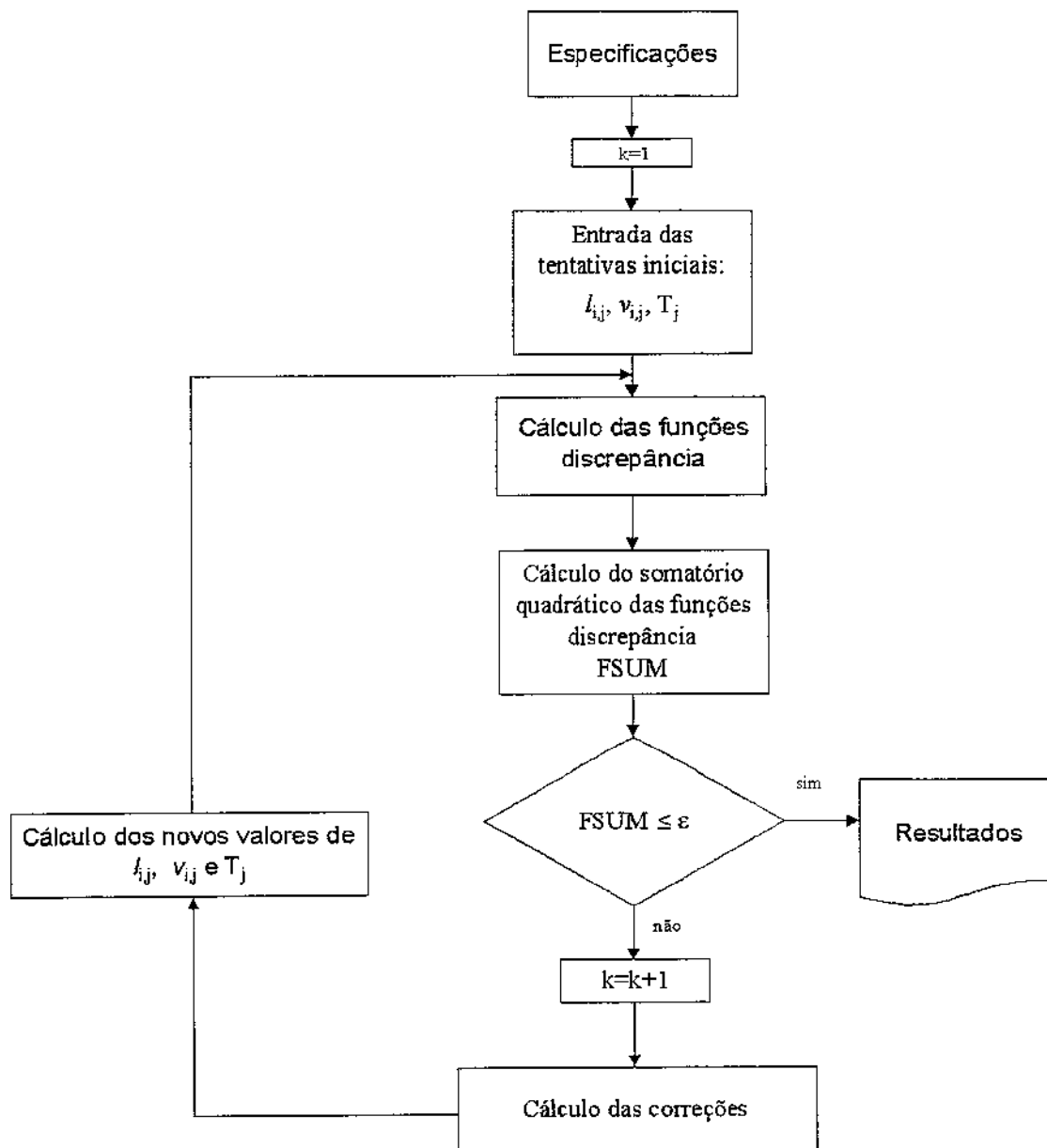


Figura A1 – Fluxograma do método de Naphtali-Sandholm.

Apêndice B – Equação de Peng-Robinson

A equação de Peng-Robinson é uma equação de estado, normalmente empregada para sistemas de hidrocarbonetos e para alguns gases inorgânicos. Sua forma generalizada é:

$$P = \frac{RT}{\bar{V} - b} - \frac{a}{\bar{V}(\bar{V} + b) + b(\bar{V} - b)}, \quad \text{C.1}$$

$$a = 0,45724 \frac{R^2 T_C^2}{P_C} \alpha, \quad \text{C.2}$$

$$b = 0,07880 \frac{RT_C}{P_C}, \quad \text{C.3}$$

$$\sqrt{\alpha} = 1 + \kappa \left(1 - \sqrt{\frac{T}{T_C}} \right), \quad \text{C.4}$$

$$\kappa = 0,37464 + 1,54226\omega - 0,26992\omega^2 \quad \text{C.5}$$

e

$$\omega = -1 - \log_{10} \left[\frac{P^{vap}(T_R = 0,7)}{P_C} \right] \quad \text{C.6}$$

onde P é a pressão, T é a temperatura, \bar{V} é o volume molar, T_C é a temperatura crítica, P_C é a pressão crítica, R é a constante dos gases, T_R é a temperatura reduzida e ω é o fator acêntrico. Os valores de a e b são dados para um componente puro. No caso de sistemas multicomponentes, estes valores são calculados a partir dos valores para componente puro,

$$a = \sum_{i=1}^{NC} \sum_{j=1}^{NC} y_i y_j a_{ij} \quad \text{C.7}$$

$$b = \sum_{i=1}^{NC} y_i b_i \quad \text{C.8}$$

onde, a_{ij} e b_i são os valores calculados para componente puro e:

$$a_{ij} = \sqrt{a_{ii} a_{jj}} (1 - k_{ij}) \quad \text{C.9}$$

O parâmetro k_{ij} é chamado de parâmetro de interação binário e pode ser encontrado em tabelas (Reid *et alli*, 1987).

A equação C.1 também pode ser escrita em função do fator de compressibilidade,

Z:

$$Z^3 - (1 - B)Z^2 + (A - 3B^2 - 2B)Z - (AB - B^2 - B^3) = 0 \quad \text{C.10}$$

onde:

$$Z = \frac{P\bar{V}}{RT}, \quad \text{C.11}$$

$$A = \frac{aP}{R^2T^2} \quad \text{C.12}$$

e

$$B = \frac{Pb}{RT} \quad \text{C.13}$$

As equações de estado podem ser usadas para o cálculo de propriedades termodinâmicas como entalpia e constante de equilíbrio. Os passos para o cálculo destas propriedades são dados a seguir:

- 1) Cálculo de b e κ ;
- 2) Cálculo de a e α para a temperatura de interesse;
- 3) Cálculo do fator de compressibilidade para as fases líquida e vapor, por meio de C.1 e C.11 ou diretamente de C.10.

Estes passos servem para o cálculo de ambas as propriedades. Para o cálculo das entalpias de líquido e vapor os seguintes procedimentos devem ser seguidos:

- Cálculo da derivada de a em relação à temperatura, por meio de:

$$\frac{da}{dT} = -0,45724 \frac{R^2T_c^2}{P_c} \kappa \sqrt{\frac{\alpha}{TT_c}} \quad \text{C.14}$$

- Cálculo das entalpias de líquido e de vapor. Através de:

$$H(T, P) = H^*(T, P) + RT(Z - 1) + \frac{T}{2\sqrt{2}b} \frac{da}{dT} - a \ln \left[\frac{Z + (1 + \sqrt{2})B}{Z + (1 - \sqrt{2})B} \right] \quad \text{C.15}$$

onde: H^* é a entalpia de referência para cada fase.

Para o cálculo da constante de equilíbrio, são seguidos os passos a seguir:

- Cálculo dos coeficientes de fugacidade para as fases líquida e vapor, por meio de:

$$\ln \hat{\phi}_i = (Z - 1) - \ln(Z - B) - \frac{A}{2\sqrt{2}B} \left[\frac{Z + (1 + \sqrt{2})B}{Z + (1 - \sqrt{2})B} \right] \quad \text{C.17}$$

onde: $\hat{\phi}_i$ é o coeficiente de fugacidade do componente i , da respectiva fase.

- Cálculo das constantes de equilíbrio, por meio de:

$$K_i = \frac{y_i}{x_i} = \frac{\hat{\phi}_i^L}{\hat{\phi}_i^V} \quad \text{C.18}$$