

RITA MARIA DE BRITO ALVES

**OTIMIZAÇÃO DE UM PROCESSO INDUSTRIAL DE
PRODUÇÃO DE ISOPRENO via REDES NEURAIIS**

**Tese apresentada à Escola Politécnica
da Universidade de São Paulo para
obtenção do título de Doutor em
Engenharia.**

São Paulo

2003

**CONSULTA
FT-1854**

OK

RITA MARIA DE BRITO ALVES

**OTIMIZAÇÃO DE UM PROCESSO INDUSTRIAL DE
PRODUÇÃO DE ISOPRENO via REDES NEURAIAS**

**Tese apresentada à Escola Politécnica da
Universidade de São Paulo para obtenção
do título de Doutor em Engenharia.**

**Área de Concentração:
Engenharia Química**

**Orientador:
Prof. Dr. Cláudio A. Oller do Nascimento**

São Paulo

2003

A *Deus*, pela vida.

**Ao meu amado marido,
Pedro.**

AGRADECIMENTOS

Ao Prof. Dr. Cláudio Augusto Oller do Nascimento pela orientação, permanente incentivo e apoio, diretrizes e discussões que permitiram o desenvolvimento deste trabalho. As idéias aqui apresentadas devem-se a ele,

À BRASKEM, na pessoa do Eng. Fernando Parente Lira Cavalcante pela liberação dos dados de processo necessários à realização deste trabalho e, em especial, à Enga. Williane Lopes Carneiro, grande amiga, cuja atenção e colaboração foram indispensáveis ao desenvolvimento e conclusão deste trabalho.

Ao meu marido, Pedro, pela paciência, companheirismo, incentivo, apoio, colaboração, compreensão e amor, que em muitos momentos foram decisivos para a finalização deste trabalho.

Aos meus pais pelo incentivo e apoio em todos os momentos da minha vida e acima de tudo por terem , desde muito cedo, me ensinado o valor da educação.

Aos colegas e amigos do Departamento de Engenharia Química pelo interesse e colaboração e, em especial, ao grupo do LSCP, que em muito contribuíram para a realização deste trabalho através do companheirismo e incentivo demonstrados ao longo destes anos.

A todos os meus amigos que compartilharam as vitórias e percalços ao longo desses anos, sempre demonstrando muito carinho, incentivo e compreensão.

À FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo pelo suporte financeiro que tornou possível a realização deste trabalho.

SUMÁRIO

RESUMO	
ABSTRACTS	
LISTA DE FIGURAS	
LISTA DE TABELAS	
NOMENCLATURA	
1. INTRODUÇÃO	1
2. REVISÃO BIBLIOGRÁFICA	
2.1. Redes Neurais	8
2.1.1. Aplicação em Engenharia Química	12
2.2. Otimização de Colunas de Destilação	15
2.3. Azeotropia	17
3. REDES NEURAS	
3.1. Introdução	21
3.2. Metodologia genérica e seqüencial de treinamento de RN	25
3.3. Aplicações em Engenharia Química	30
3.4. Estrutura Matemática da Rede e o Algoritmo de Aprendizado	30
4. OTIMIZAÇÃO QUALITATIVA/QUANTITATIVA DE PROCESSOS INDUSTRIAIS	
4.1. Introdução	36
4.2. Modelagem e Simulação do Processo	
4.2.1. Metodologia	40
4.2.2. Modelagem e Ajuste do Sistema	40
4.3. Otimização de Processos	42
4.3.1. Algoritmo de Otimização	44
4.3.2. Abordagem para Otimização com Base em Redes Neurais	45
4.3.3. Metodologia	46
4.4. Aplicação à Planta de Isopreno	
4.4.1. Descrição do Processo	50
4.4.2. Metodologia	59

4.4.2.1.	Dados Industriais	61
4.4.2.2.	Tratamento dos Dados	61
4.4.2.3.	Modelagem	66
4.4.2.4.	Otimização	68
4.5.	Resultados e Discussão	
4.5.1.	Preparação dos Dados	69
4.5.2.	Modelagem	80
4.5.2.1.	Subsistema 1 – REDE 1	80
4.5.2.2.	Subsistema 2 – REDE 2	86
4.5.2.3.	Subsistema 3 – REDE 3	91
4.5.2.4.	Subsistema 4 – REDE 4	95
4.5.2.5.	Subsistema 5 – REDE 5	100
4.5.2.6.	Subsistema 6 – REDE 6	105
4.5.2.7.	Subsistema 7 – REDE 7	110
4.5.2.8.	Subsistema 8 – REDE 8	114
4.5.2.9.	Subsistema 9 – REDE 9	118
4.5.2.10.	Subsistema 10 – REDE 10	123
4.5.3.	Modelagem Global do Processo	127
4.5.4.	Simulação do Processo Global	138
4.5.5.	Análise de Sensibilidade do Processo	139
4.5.5.1.	Reator – R-01	140
4.5.5.2.	Primeira Coluna de Remoção de Pesados – COL-01	152
4.5.6.	Otimização do Processo	156
4.6.	Conclusões	167
5.	PREDIÇÃO DE AZEÓTROPOS	
5.1.	Introdução	169
5.2.	Princípios Básicos da Azeotropia	170
5.3.	Metodologia	177
5.3.1.	Seleção das Variáveis da Rede	179
5.3.2.	Banco de Dados de Sistemas Binários Azeotrópicos	181
5.3.3.	Predição da Formação de Azeótropos	

5.3.3.1.	Seleção e Treinamento da Rede Neural	182
5.3.3.2.	Validação da Rede Neural e Análise dos Resultados	193
5.3.4.	Predição das Condições do Ponto Azeotrópico	195
5.4.	Conclusões	198
6.	CONCLUSÕES	199
	ANEXO A – Seleção do Número de Neurônios das Redes Neurais	201
	ANEXO B – Ajuste das Redes Neurais – Gráficos	207
	ANEXO C – Simulação do Processo – Tabelas	239
	ANEXO D – Análise de Sensibilidade – Tabelas	250
	ANEXO E – Resultados da Otimização do Processo – Tabelas	255
	ANEXO F – Resultado de PCA para Azeotropia	268
	LISTA DE REFERÊNCIAS	271
	APÊNDICE A – Artigos	

RESUMO

Este trabalho descreve a aplicação de redes neurais “feed-forward” com três camadas em diferentes áreas da Engenharia Química. O objetivo principal do projeto é a modelagem, simulação e posterior otimização do processo de produção de isopreno empregando técnicas de redes neurais em substituição as equações de modelagem fenomenológica. A planta industrial testada é a unidade de produção de isopreno da BRASKEM (antiga COPENE). O sistema consiste essencialmente de um reator de dimerização e uma série de colunas de destilação. Uma vez que redes neurais são capazes de aprender eficientemente o processo a partir de informações extraídas diretamente de dados da planta, para este trabalho o modelo de rede neural gerado foi construído a partir de dados históricos operacionais coletados a cada 15 minutos durante o período de 1 ano. Em uma primeira etapa é realizada a análise dos dados operacionais de modo a detectar e eliminar erros grosseiros e sistemáticos. Em seguida, a modelagem e simulação do processo são realizadas. O modelo de redes neurais gerado é, então, empregado na otimização qualitativa/quantitativa do processo, construindo um “grid” de busca detalhado da região de interesse, através um mapeamento completo da função objetivo no espaço das variáveis de decisão. A segunda etapa diz respeito à predição de azeótropos, visando um melhor entendimento do comportamento do sistema da seção de extração de isopreno. Nas duas etapas, a grande vantagem em utilizar modelos de redes neurais, além de ajustar dados, é a capacidade que estes apresentam em representar eficientemente sistemas multivariáveis, complexos e não lineares, aprendendo o sistema, sem o conhecimento das leis físicas e químicas que o regem. Comparações entre a predição dos modelos propostos e os dados experimentais foram executadas e resultados muito bons foram conseguidos do ponto de vista industrial. Esta metodologia fornece informações interessantes e de maior compreensão para a análise dos engenheiros de processo do que os procedimentos convencionais correspondentes. Além disso, este trabalho mostra que a metodologia de redes neurais é promissora para varias aplicações industriais, tais como análise de dados, modelagem, simulação e otimização de processos, bem como predição de propriedades termodinâmicas.

ABSTRACT

This work describes the application of a three-layer feed-forward neural network (NN) in different areas of chemical engineering. The main objective of this study is to model, simulate and optimize a real industrial plant, using NN by replacing phenomenological models. The industrial process studied is the isoprene production unit from BRASKEM. The chemical process consists basically of a dimerization reactor and a separation column train. Since NNs are able to extract information from plant data in an efficient manner, for this work, the neural network model was built directly from historical plant data, which were collected every 15 minutes during a period of one year. These data were carefully analyzed in order to identify and eliminate gross error data and non-steady state operation data. The modeling using NN was carried out by parts in order to get information on intermediate streams. Then, the global model was built, by interconnecting each individual model, and used to simulate and optimize the process. The optimization procedure carries on a detailed grid search of the region of interest, by a full mapping of the objective function on the space of decision variables. The second stage of this work deals with the azeotropic prediction using also the neural network approach. The objective of this step was to obtain a better understanding of the system behavior in the isoprene extraction section. Since all the cases studied are non-linear, complex and multivariable systems, NN approach appears as a technique of interest due to its capability of learning the system without knowledge of the physical and chemical laws that govern it. Comparisons between the model's prediction and the experimental data were performed and reasonable results were achieved from an industrial point of view. Using neural network approach provides more comprehensive information for an engineer's analysis than the conventional procedure. This work shows that the use of NN methodology is promising for several industrial applications, such as data analysis, modeling, simulation and optimization process, as well as thermodynamics properties prediction. However, success in obtaining a reliable and robust NN depends strongly on the choice of the variables involved, as well as the quality of available data set and the domain used for training purposes.

LISTA DE FIGURAS

Figura 3.1.	Rede de Múltiplas Camadas	22
Figura 3.2.	Rede Recursiva	22
Figura 3.3.	Representação de um neurônio da camada j	31
Figura 4.1.	Esquema do Modelo da Planta Industrial – “Caixa Preta”	42
Figura 4.2.	Fluxograma de Processo da Unidade de Isopreno	51
Figura 4.3.	Esquema do Procedimento de Otimização	60
Figura 4.4.	Erro Mínimo para o Conjunto de Teste – Exemplo	67
Figura 4.5.	Conjunto de Dados Experimentais Originais	70
Figura 4.6.	Conjunto de Dados Experimentais Após Eliminação de “outliers”	70
Figura 4.7.	Conjunto de Dados Experimentais Após Eliminação de “outliers” e erros sistemáticos	71
Figura 4.8.	Balanço Material Global Antes do Ajuste e Eliminação de Erros Sistemáticos	72
Figura 4.9.	Balanço Material Global Após Ajuste e Eliminação de Erros Sistemáticos	73
Figura 4.10.	Histograma do Balanço de IP Antes da Correção e Eliminação de Erros Sistemáticos	74
Figura 4.11.	Histograma do Balanço de IP após correção e eliminação de erros Sistemáticos	74
Figura 4.12.a.	Comparação entre valores calculados e dados medidos: antes da análise	78
Figura 4.12.b.	Comparação entre valores calculados e dados medidos: resultados intermediários	78
Figura 4.12.c.	Comparação entre valores calculados e dados medidos: resultado final	79
Figura 4.13.	Fluxograma do reator modelado pela REDE 1	81
Figura 4.14.	Esquema da REDE 1 – variáveis de entrada/saída	81
Figura 4.15.	Fluxograma da COL-01 modelada pela REDE 2	86

Figura 4.16.	Esquema da REDE 2 – variáveis de entrada/saída	87
Figura 4.17.	Fluxograma da COL-12 modelada pela REDE 3	91
Figura 4.18.	Esquema da REDE 3 – variáveis de entrada/saída	92
Figura 4.19.	Fluxograma da COL-02 modelada pela REDE 4	95
Figura 4.20.	Esquema da REDE 4 – variáveis de entrada/saída	96
Figura 4.21.	Fluxograma das COL-03 e COL-04 modeladas pela REDE 5	100
Figura 4.22.	Esquema da REDE 5 – variáveis de entrada/saída	101
Figura 4.23.	Fluxograma da COL-05 modelada pela REDE 6	105
Figura 4.24.	Esquema da REDE 6 – variáveis de entrada/saída	106
Figura 4.25.	Fluxograma das COL-06 e COL-07 modeladas pela REDE 7	110
Figura 4.26.	Esquema da REDE 7 – variáveis de entrada/saída	111
Figura 4.27.	Fluxograma da COL-10 modelada pela REDE 8	114
Figura 4.28.	Esquema da REDE 8 – variáveis de entrada/saída	115
Figura 4.29.	Fluxograma da COL-08 modelada pela REDE 9	118
Figura 4.30.	Esquema da REDE 9 – variáveis de entrada/saída	119
Figura 4.31.	Fluxograma da COL-09 modelada pela REDE 10	123
Figura 4.32.	Esquema da REDE 10 – variáveis de entrada/saída	124
Figura 4.33.	Representação de Fluxo da Seção de Preparação de Carga	129
Figura 4.34.	Sistema de Extração - Primeira Coluna de Destilação Extrativa	130
Figura 4.35.	Sistema de Extração – Segunda Coluna de Destilação Extrativa e Stripper de Solvente	131
Figura 4.36.	Sistema de Lavagem e Recuperação de Solvente	132
Figura 4.37.	Seção de Fracionamento	133
Figura 4.38.	Esquema de Modelagem usando Redes Neurais - Simulação do Processo Global	134
Figura 4.39.	Influência da Temperatura da Carga do Reator na Conversão de CPD – Caso 1	142
Figura 4.40.	Influência da Temperatura da Carga do Reator na Perda de IP no Reator – Caso 1	143
Figura 4.41.	Influência da Temperatura da Carga do Reator na Conversão de CPD – Caso 2	144

Figura 4.42.	Influência da Temperatura da Carga do Reator na Perda de IP no Reator – Caso 2	145
Figura 4.43.	Influência da Temperatura da Carga do Reator na Conversão de CPD – Caso 3	146
Figura 4.44.	Influência da Temperatura da Carga do Reator na Perda de IP no Reator – Caso 3	147
Figura 4.45.	Influência da Temperatura da Carga do Reator na Conversão de CPD – Caso 4	149
Figura 4.46.	Influência da Temperatura da Carga do Reator na perda de IP no Reator – Caso 4	150
Figura 5.1.	Exemplos Típicos de Misturas Azeotrópicas	174
Figura 5.2.	Representação da Rede Neural – Predição de Azeótropos	178
Figura 5.3.	Representação Esquemática da Rede Neural (RN1)- Predição de azeótropos	183
Figura 5.4.	Resultado da Rede Neural – Predição de Azeótropos - Caso I	184
Figura 5.5.	Resultado da Rede Neural – Predição de Azeótropos - Caso II	187
Figura 5.6.	Resultado da Rede Neural – Predição de Azeótropos - Caso III	188
Figura 5.7.	Resultado da Rede Neural – Predição de Azeótropos - Caso IV	190
Figura 5.8.	Resultado da Rede Neural – Predição de Azeótropos - Caso V	190
Figura 5.9.	Representação Esquemática da Rede Neural (RN2)- Predição das condições azeotrópicas	196
Figuras 5.10.a-b.	Predição das condições azeotrópicas a 1 atm Comparação entre Redes Neurais e UNIFAC	197
Figura A.1.	Seleção do Número de Neurônio da REDE 1	202
Figura A.2.	Seleção do Número de Neurônio da REDE 2	202
Figura A.3.	Seleção do Número de Neurônio da REDE 3	203
Figura A.4.	Seleção do Número de Neurônio da REDE 4	203
Figura A.5.	Seleção do Número de Neurônio da REDE 5	204
Figura A.6.	Seleção do Número de Neurônio da REDE 6	204
Figura A.7.	Seleção do Número de Neurônio da REDE 7	205
Figura A.8.	Seleção do Número de Neurônio da REDE 8	205

Figura A.9.	Seleção do Número de Neurônio da REDE 9	206
Figura A.10.	Seleção do Número de Neurônio da REDE 10	206
Figura B.1.	Ajuste da REDE 1 – R-01 para o Conjunto de Teste	208
Figura B.2.	Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 1 – R-01	208
Figura B.3.	Histograma do Erro – Variáveis de Saída da REDE 1- R-01	209
Figura B.4.	Ajuste da REDE 2 – COL-01 para o Conjunto de Teste	210
Figura B.5.	Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 2 – COL-01	211
Figura B.6.	Histograma do Erro – Variáveis de Saída da REDE 2- COL-01	212
Figura B.7.	Ajuste da REDE 3 – COL-12 para o Conjunto de Teste	213
Figura B.8.	Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 3 – COL-12	214
Figura B.9.	Histograma do Erro – Variáveis de Saída da REDE 3- COL-12	215
Figura B.10.	Ajuste da REDE 4 – COL-02 para o Conjunto de Teste	216
Figura B.11.	Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 4 – COL-02	217
Figura B.12.	Histograma do Erro – Variáveis de Saída da REDE 4- COL-02	218
Figura B.13.	Ajuste da REDE 5 – COL-03 e COL-04 para o Conjunto de Teste	219
Figura B.14.	Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 5 – COL-03 e COL-04	221
Figura B.15.	Histograma do Erro – Variáveis de Saída da REDE 5 – COL-03 e COL-04	223
Figura B.16.	Ajuste da REDE 6 – COL-05 para o Conjunto de Teste	225
Figura B.17.	Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 6 – COL-05	225
Figura B.18.	Histograma do Erro – Variáveis de Saída da REDE 6 – COL-05	226
Figura B.19.	Ajuste da REDE 7 – COL-06 e COL-07 para o Conjunto de Teste	227

Figura B.20. Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 7 – COL-06 e COL-07	228
Figura B.21. Histograma do Erro – Variáveis de Saída da REDE 7 – COL-06 e COL-07	229
Figura B.22. Ajuste da REDE 8 – COL-10 para o Conjunto de Teste	230
Figura B.23. Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 8 – COL-10	231
Figura B.24. Histograma do Erro – Variáveis de Saída da REDE 8 – COL-10	232
Figura B.25. Ajuste da REDE 9 – COL-08 para o Conjunto de Teste	233
Figura B.26. Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 9 – COL-08	234
Figura B.27. Histograma do Erro – Variáveis de Saída da REDE 9 – COL-08	235
Figura B.28. Ajuste da REDE 10 – COL-09 para o Conjunto de Teste	236
Figura B.29. Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 10 – COL-09	237
Figura B.30. Histograma do Erro – Variáveis de Saída da REDE 10 – COL-09	238

LISTA DE TABELAS

Tabela 4.1.	Análise de Grupos – Exemplos	75
Tabela 4.2.	Análise de Teste de Hipótese das Médias – Exemplos	76
Tabela 4.3.	Pontos eliminados em cada corrida – Exemplo	77
Tabela 4.4.	Variáveis de Processo para a REDE 1 e Suas Faixas de Aplicação	82
Tabela 4.5.	Configurações da REDE 1	83
Tabela 4.6.	Variáveis de Processo para a REDE 2 e Suas Faixas de Aplicação	87
Tabela 4.7.	Configurações da REDE 2	89
Tabela 4.8.	Variáveis de Processo para a REDE 3 e Suas Faixas de Aplicação	92
Tabela 4.9.	Configurações da REDE 3	93
Tabela 4.10.	Variáveis de Processo para a REDE 4 e Suas Faixas de Aplicação	96
Tabela 4.11.	Configurações da REDE 4	98
Tabela 4.12.	Variáveis de Processo para a REDE 5 e Suas Faixas de Aplicação	101
Tabela 4.13.	Configurações da REDE 5	103
Tabela 4.14.	Variáveis de Processo para a REDE 6 e Suas Faixas de Aplicação	106
Tabela 4.15.	Configurações da REDE 6	107
Tabela 4.16.	Variáveis de Processo para a REDE 7 e Suas Faixas de Aplicação	110
Tabela 4.17.	Configurações da REDE 7	112
Tabela 4.18.	Variáveis de Processo para a REDE 8 e Suas Faixas de Aplicação	115
Tabela 4.19.	Configurações da REDE 8	116
Tabela 4.20.	Variáveis de Processo para a REDE 9 e Suas Faixas de Aplicação	119

Tabela 4.21.	Configurações da REDE 9	121
Tabela 4.22.	Variáveis de Processo para a REDE 10 e Suas Faixas de Aplicação	124
Tabela 4.23.	Configurações da REDE 10	126
Tabela 4.24.	Variáveis Modeladas para o Processo de Produção de Isopreno	135
Tabela 4.25.	Caso Base para Análise de Sensibilidade	141
Tabela 4.26.	Influência do Aumento da Vazão de Refluxo na Vazão e Composição do Topo da COL-01	154
Tabela 4.27.	Variáveis Manipuladas e Restrições para a Otimização do Processo	157
Tabela 4.28.	Variáveis de Entrada para o Caso Base para Otimização	160
Tabela 4.29.	Otimização Global – Condição 1	163
Tabela 4.30.	Otimização Global – Condição 2	164
Tabela 4.31.	Otimização Global – Condição 3	165
Tabela 4.32.	Otimização Global – Condição 4	166
Tabela 5.1.	Desvios da Lei de Raoult	173
Tabela 5.2.	Comportamento Azeotrópico Discrepante – Exemplo	182
Tabela 5.3.	Comparação entre Valores Experimentais e Preditos Utilizando Modelo de Redes Neurais e Método UNIFAC – TESTE	192
Tabela 5.4.	Comparação entre Valores Experimentais e Preditos Utilizando Modelo de Redes Neurais e Método UNIFAC – VALID	194
Tabela 5.5.	Dados Azeotrópicos Experimentais para Sistemas Formados por Álcoois e Cetonas	195
Tabela C.1.	Valores das Variáveis de Entrada para Simulação	240
Tabela C.2.	Resultados da Simulação do Processo Global	242
Tabela D.1.	Teor de CPD no Reator – $y_1(2)$ – Caso 1	251
Tabela D.2.	Perda de IP no Reator – $x_1(2)$ - $y_1(1)$ – Caso 1	251
Tabela D.3.	Teor de CPD no Reator – $y_1(2)$ – Caso 2	252
Tabela D.4.	Perda de IP no Reator – $x_1(2)$ - $y_1(1)$ – Caso 2	252
Tabela D.5.	Teor de CPD no Reator – $y_1(2)$ – Caso 3	253
Tabela D.6.	Perda de IP no Reator – $x_1(2)$ - $y_1(1)$ – Caso 3	253

Tabela D.7.	Teor de CPD no Reator – $y_1(2)$ – Caso 4	254
Tabela D.8.	Perda de IP no Reator – $x_1(2) - y_1(1)$ – Caso 4	254
Tabela E.1.	Resultados da Otimização da Seção de Preparação de Carga	256
Tabela E.2.	Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 1 Ótimo da Seção de Preparação de Carga	257
Tabela E.3.	Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 2 Ótimo da Seção de Preparação de Carga	259
Tabela E.4.	Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 3 Ótimo da Seção de Preparação de Carga	261
Tabela E.5.	Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 4 Ótimo da Seção de Preparação de Carga	263
Tabela E.6.	Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 5 Ótimo da Seção de Preparação de Carga	265
Tabela E.7.	Resultados da Otimização da Seção de Fracionamento	267

NOMENCLATURA

AAL	Álcool Alílico
ACN	Acetonitrila
C10	Hidrocarbonetos com 10 átomos de Carbono
C5 ou C ₅	Hidrocarbonetos com 5 átomos de Carbono
n-C5 ou n-C ₅	n-Pentano
i-C5 ou i-C ₅	iso-Pentano
CPD	Ciclopentadieno
DCPD	Diciclopentadieno
H2O ou H ₂ O	Água
IP	Isopreno
2M2B	2-Metil-2-Buteno
PP	Piperileno
1,4PTD	1,4-Pentadieno
NH	Número de Neurônios na Camada Oculta de uma Rede Neural
RN	Rede Neural
#	em uma coluna de destilação significa: prato (estágio) número . ex.: # 10 - prato 10 da coluna
Δ	Delta - Variação

1. INTRODUÇÃO

Uma característica peculiar dos processos químicos é que a maioria deles apresenta forte não linearidade devido aos inúmeros fenômenos físicos e químicos associados ao mesmo tempo, o que dificulta e, por vezes, impede a obtenção de modelos fenomenológicos, a ponto de, em alguns casos, conseguir-se apenas um modelo extremamente simplificado e válido somente em condições muito específicas. É exatamente nessa área que as redes neurais têm apresentado grande potencial de aplicabilidade, devido a sua capacidade de “aprender” o que ocorre no processo, sem conhecimento das leis físicas e químicas que o regem.

Este trabalho descreve a aplicação de um modelo de redes neurais (RN) ao processo industrial para produção de Isopreno. O objetivo desta planta é produzir isopreno de alta pureza para obtenção de borracha sintética e termoplástica a partir de uma corrente C_5 proveniente da unidade de gasolina de pirólise. Este processo é complexo e de difícil simulação empregando ferramentas convencionais, principalmente devido à falta de dados termodinâmicos (dados de equilíbrio) que representem bem o sistema.

O trabalho foi dividido em duas partes, ambas empregando Redes Neurais como ferramenta para correlacionar as informações pertinentes a cada caso, exemplificando a aplicabilidade das mesmas em diferentes áreas da Engenharia Química.

1.1. Otimização qualitativa/quantitativa da unidade de Isopreno:

O processo de produção de Isopreno pode ser dividido em cinco seções: preparação da carga, destilação extrativa e recuperação de solvente, remoção de enxofre e fracionamento. O sistema consiste basicamente de um reator de dimerização e uma série de colunas de destilação.

A idéia básica é a modelagem e posterior otimização deste processo petroquímico usando a metodologia de redes neurais para correlacionar dados históricos de operação da unidade, em substituição aos modelos fenomenológicos. Entretanto, o sucesso na obtenção de uma rede neural robusta e apropriada depende fortemente da escolha das variáveis de processo envolvidas, bem como da qualidade dos dados disponíveis e do domínio usado para fins de treinamento. Se qualquer comportamento importante do processo não for coberto por observações, o modelo de redes neurais ou qualquer outro modelo empírico terá uma representação mais pobre e a otimização provavelmente falhará

Usar Redes Neurais, para resolver problemas de otimização, implica no levantamento total da superfície de resposta do processo com restrições. A partir dessa superfície pode-se chegar aos mínimos globais das variáveis ou até mesmo determinar a região ótima de operação. Alternativas recentes para procedimentos de otimização baseada no uso de redes neurais torna possível uma busca detalhada do “grid search”. A idéia principal é substituir as equações de modelagem por uma rede neural equivalente, empregando dados históricos da planta industrial e utilizar esta rede para executar um “grid” de busca detalhado da região de interesse, através de um mapeamento completo da função objetivo no espaço das variáveis de decisão. Usar redes neurais em substituição ao modelo fenomenológico apresenta a vantagem de resposta mais rápida da simulação via redes neurais. Desta forma, o “grid” detalhado pode ser atingido em tempo razoável, desde que não existam muitas variáveis sendo otimizadas, o que levaria a possíveis problemas de dimensionalidade. Esta abordagem é mais confiável, lida com restrições, evita vários problemas numéricos típicos das ferramentas convencionais de otimização e apresenta maior velocidade computacional. Como benefício adicional, o mapeamento completo da função objetivo permite identificar facilmente múltiplos pontos ótimos, uma importante característica não apresentada pelos métodos convencionais de otimização. Além disso, as restrições são facilmente tratadas desde que pontos com restrições violadas possam ser reconhecidos e classificados de acordo com a importância da restrição. Uma vez que o mapa é obtido, é fácil escolher o ponto

ótimo, identificar se múltiplos ótimos estão presentes, checar se restrições foram violadas, entre outros.

Esta abordagem definitivamente fornece mais informações para análise do engenheiro que os procedimentos convencionais de programação não linear. Este algoritmo pode ser estendido diretamente para tratar problemas de otimização multi-objetiva.

No caso em estudo, o modelo foi desenvolvido via redes neurais devido principalmente a grande dificuldade na modelagem fenomenológica da Seção de Extração. Assim, a otimização do comportamento das variáveis deste processo é possível utilizando experiências passadas e informações adicionais como restrições, sem a necessidade do conhecimento físico-químico do processo. No caso estudado, inicialmente trabalhou-se com dados históricos formando um banco de dados com valores médios diários da planta englobando três anos de produção, o que trouxe alguns problemas no ajuste das redes neurais, por motivos de falta confiabilidade dos dados. Numa etapa posterior, foram usados dados obtidos com frequência de 15 minutos durante o período de 01 ano, o que facilitou o estabelecimento de períodos de estado estacionário e possibilitou avaliar problemas nos instrumentos de medição. Assim, o banco de dados original usado é constituído por 34500 conjuntos de observações de 244 variáveis.

Estes dados foram cuidadosamente analisados de modo a identificar e eliminar erros grosseiros e sistemáticos e períodos em regime de operação não estacionário. Quando se utilizam dados obtidos diretamente da planta industrial, é necessário tratar tais dados de modo a detectar erros grosseiros e/ou sistemáticos. Esta fase do trabalho é de extrema importância já que a qualidade dos dados influi diretamente na qualidade do ajuste realizado pelas redes neurais.

Em alguns casos, “outliers” podem ser facilmente detectados, mas para outros não é tão óbvio. Se a origem dos valores anômalos é conhecida, pode-se sem hesitação descartar tais observações. Por outro lado, quando um erro ou ~~uma~~

observação extrema não é claramente justificável, o julgamento deve ser baseado em algum tipo de análise estatística. Neste trabalho, a metodologia empregada envolve uma aproximação baseada em procedimentos estatísticos e outros baseados nas equações do primeiro princípio ou a combinação de ambos, além do conhecimento do processo. Além disso, o próprio modelo de redes neurais foi utilizado para representar o processo com treinamento em poucas iterações. Então os pontos nos quais os erros entre os dados experimentais e os valores calculados pela rede aparecem muito dispersos da maioria dos valores são provavelmente “outliers” e são analisados classificando entradas e saídas similares de modo a identificar “clusters”. Então, é executada a avaliação dos erros através do princípio da similaridade ou teste de hipótese de médias.

Com os dados de entrada e saída previamente selecionados, pode-se modelar as redes neurais sob diversas óticas, dentre elas: redes que modelam o processo como um todo ou redes que modelam partes do processo individualmente e através de conexão posterior entre elas, modelam o sistema completo. No caso em estudo, optou-se por dividir o processo em subsistemas. Cada subsistema representa uma grande operação unitária (reator, coluna de destilação ou extração). Assim, para cada subsistema são analisadas as variáveis medidas de entrada e saída do processo com respeito à consistência da informação e sua importância na operação do processo em estudo. Entre os ganhos de informação que a etapa de modelagem das redes traz, está a detecção de erros de medição, indicando qual provável instrumento apresentou problemas em dado período.

Após a modelagem das redes neurais individuais que representam os subsistemas, o foco principal é a interligação seqüencial dessas redes com a finalidade de gerar dados para construir a superfície de resposta das variáveis envolvidas, assim como analisar o comportamento das principais variáveis do processo. O procedimento de convergência do sistema integrado é realizado interligando as correntes comuns aos equipamentos, reproduzindo a unidade industrial com todas as correntes modeladas

pelas redes neurais. Algoritmo de substituição direta com fator de amortecimento foi utilizado para a convergência matemática.

Uma vez treinada, a rede neural pode ser usada para simulação do processo com grande rapidez. Milhares de simulações podem ser processadas em segundos, o que possibilita o emprego do modelo de redes neurais em procedimentos de simulação, otimização, controle de processo em linha, análise de restrições do processo, análise de sensibilidade das variáveis, analisadores virtuais, entre outras aplicações.

1.2. Predição de Azeótropos

O isopreno é produzido na BRASKEM por um processo de destilação extrativa que utiliza como solvente uma mistura de acetonitrila, água, álcool alílico e acetona. Dentre estes compostos, a água é o que apresenta maior oscilação na composição devido a diversos fatores como: formação de azeótropo com a acetonitrila, existência no circuito de solvente de diversos pontos de saídas e entradas de correntes com teores de água diferentes, vazamento de água para o processo pelo sistema de selagem das bombas, etc.

A estabilidade do teor de água no solvente circulante ainda é um desafio operacional para a unidade. Conteúdo de água na alimentação do solvente muito elevado pode causar separação de fases e operação instável das extratoras (inundação da coluna no pior caso). Valores muito baixos podem causar diminuição da eficiência de separação, aumento de temperatura de fundo da primeira coluna de destilação extrativa. A seção de extração de isopreno, usando a mistura de acetonitrila, água, álcool alílico e acetona, provoca, em diversos pontos do processo, o desvio da não idealidade da fase líquida com a formação de diversos azeótropos entre estes compostos e destes com diversos hidrocarbonetos. Este fato merece uma maior atenção no estudo da unidade de modo a melhor entender estes efeitos no binário acetonitrila - água, compostos predominantes no solvente circulante.

Desta forma, um maior conhecimento sobre a formação de azeótropos torna-se necessário para melhor compreensão dos princípios teóricos envolvidos. Neste trabalho, partindo-se de uma revisão ampla sobre a teoria de formação de azeótropos, chegou-se a um procedimento para predição azeotrópica entre componentes em uma mistura binária e o correspondente ponto azeotrópico, sem executar cálculos de equilíbrio líquido – vapor. O método correlaciona dados binários azeotrópicos com propriedades macroscópicas e moleculares dos componentes puros constituintes da mistura. A metodologia aplicada faz uso de redes neurais para correlacionar os sistemas. As propriedades macroscópicas examinadas incluem massa molecular, ponto de ebulição, constantes críticas e fator acêntrico, enquanto as propriedades moleculares consistem de parâmetros relacionados a dipolaridade molecular e a capacidade de cada componente em formar pontes de hidrogênio.

A vantagem de usar redes neurais, além de sua capacidade em ajustar dados, é não ser necessário o emprego de modelos termodinâmicos que representem satisfatoriamente a não idealidade da fase líquida.

Assim, este trabalho mostra, com sucesso, a diversidade de aplicações de modelos de redes neurais em problemas de engenharia química, tais como: análise de dados, modelagem, simulação e otimização de processos, e predição de propriedades termodinâmicas.

A apresentação do trabalho é dividida em capítulos, a saber:

No capítulo 2 é apresentada uma revisão bibliográfica sobre os temas abordados neste trabalho: Redes Neurais, Azeotropia e Otimização de Processos.

No capítulo 3 são mostrados aspectos referentes à fundamentação teórica relacionada ao desenvolvimento de modelos de Redes Neurais, sua estrutura e algoritmo de treinamento.

No capítulo 4 é apresentada uma aplicação de modelos de Redes Neurais na Modelagem, Simulação e Otimização do Processo Industrial de Isopreno, a partir de dados históricos da planta, o que envolve a necessidade de análise dos dados, realizada, em parte neste trabalho, empregando redes neurais, o que também é descrito em detalhes neste capítulo.

O capítulo 5 apresenta a aplicação de modelos de Redes Neurais na Predição de Propriedades Termodinâmicas, mais precisamente na predição da capacidade de uma dada mistura binária formar ou não um azeótropo e as condições azeotrópicas correspondentes (tipo, temperatura e composição) à pressão de 1 atm. São descritas neste capítulo as bases físicas fundamentais necessárias para a construção do modelo de predição proposto. São também apresentados os resultados e discussões em relação à comparação com o modelo UNIFAC, tradicionalmente utilizado.

O capítulo 6 apresenta a conclusão geral do trabalho quanto à aplicabilidade de modelos de redes neurais na área de Engenharia Química. Neste capítulo também são apresentadas sugestões para trabalhos posteriores.

Neste trabalho, para treinamento da rede foi utilizado o programa NEURO14, desenvolvido no Laboratório de Simulação e Controle de Processos do Departamento de Engenharia Química/Escola Politécnica da USP (Nascimento, 1992).

2. REVISÃO BIBLIOGRÁFICA

2.1. Redes Neurais

Rede Neural Artificial é freqüentemente identificada como uma subespecialidade da Inteligência Artificial. Foi muito utilizada para problemas de classificação e reconhecimento de padrões, como parte da teoria conexionista dos processos mentais e também como uma categoria de modelos em ciência da cognição.

A idéia de utilizar redes neurais artificiais como estratégia de solução de problemas que requerem análises complexas de dados não é recente. O campo de cálculo neural é quase tão antigo quanto à computação digital propriamente dita. Pesquisa em rede neural pode ser referenciada aos anos 1940, quando trabalhos realizados por McCulloch e Pitts (1943) sobre modelagem de neurônios e por Hebb (1949) sobre aprendizagem foram publicados. Mais especificamente, em 1943, inspirado talvez pelos avanços na determinação experimental das propriedades eletrofisiológicas do neurônio, e pelos então recentes resultados de Alan Turing e John von Neumann que indicavam ser booleana a natureza essencial da inteligência, o que por sua vez resultava possivelmente da atividade neuronal, Warren McCulloch, lembrado nos anais da ciência como médico, filósofo, matemático e poeta, juntamente com o estatístico Walter Pitts publicaram no *Bulletin of Mathematical Biophysics* um artigo intitulado “*A Logical Calculus of the Ideas Immanent in Nervous Activity*” que hoje é a referência número 1 para a teoria das redes neurais artificiais (Kóvacs, 1996).

A essência da proposta de McCulloch e Pitts foi a seguinte: A inteligência é equivalente ao cálculo de predicados, que por sua vez podem ser implementados por funções booleanas. Por outro lado, o sistema nervoso é composto de redes de neurônios, que, com as devidas simplificações, têm a capacidade básica de implementar estas funções booleanas. Conclusão: a ligação entre inteligência e atividade nervosa fica estabelecida de forma científica (Kóvacs, 1996).

No final da década de 1950, Rosenblatt (1958), na Universidade de Cornell, deu prosseguimento às idéias de McCulloch. Ele criou uma genuína rede de múltiplos neurônios do tipo discriminadores lineares e chamou esta rede de *perceptron*. Um perceptron é uma rede com os neurônios dispostos em várias camadas. O perceptron foi capaz de aprender a calcular algumas funções lógicas modificando as forças das conexões sinápticas.

Na mesma época que Rosenblatt trabalhava no perceptron, Widrow, na Universidade de Stanford, desenvolveu um modelo neural linear, muito simples conceitualmente, que ele batizou de ADALINE (ADaptive LINEar Element) e mais tarde sua generalização multidimensional, o MADALINE (Múltipla ADALINE). A contribuição realmente importante do trabalho de Widrow foi a invenção de um princípio de treinamento extremamente poderoso para as redes de adalines conhecido como *Regra Delta*, que foi mais tarde generalizado para redes com modelos neurais mais elaborados (Widrow, 1962). A Regra Delta ou método do gradiente propõe obter o ponto de mínimo através de um processo de iteração local, utilizando um exemplo do conjunto de treinamento por vez. O método do gradiente aplicado ao problema de determinação da melhor aproximação linear a um conjunto de dados não passaria de um exercício acadêmico de otimização em espaços lineares, não tivesse preparado o terreno para a introdução do mais poderoso procedimento que se conhece até hoje para treinamento de redes neurais, o método de retropropagação do erro, desenvolvido principalmente por Rummelhart, Hint e Williams (1986),

Em 1962, entram em cena Marvin Minsky e Seymour Pappert com o agora clássico texto *Perceptrons* (Minsky & Pappert, 1969). Neste livro, reeditado em 1991, os autores argumentam enfaticamente quanto às limitações básicas de perceptrons isolados.

Os anais da história de redes neurais anotam a publicação de *Perceptrons* como um divisor de águas que paralisou a pesquisa nesta área na década de 1970. Segundo Rummelhart et al. (1986), deu-se então o início da era do obscurantismo em pesquisa com redes neurais que só foi encerrado no início da década de 1980 com o trabalho

de Hopfield em memórias associativas (Hopfield, 1982). No entanto, a área de redes neurais teve um desenvolvimento explosivo após a publicação, em 1986, do hoje clássico “Parallel Distributed Processing”, editado por Rumelhart e McClelland do PDP Research Group da Universidade da Califórnia, em São Diego.

Um breve histórico de redes neurais é apresentado por Nelson e Illingworth (1991), mostrando algumas de suas facetas interessantes. Outra referência geral sobre redes neurais pode ser vista em Dayhoff (1990). Kóvacs (1996) também apresenta um resumo sobre redes neurais e seu treinamento. Haykin(1994) faz uma apresentação mais detalhada e extensa sobre redes neurais. Mukesh (1997) apresenta uma analogia didática entre redes neurais artificiais e o cérebro humano e exemplifica a aplicação de rede neural como uma alternativa para regressão matemática de conjunto complexo de dados.

Redes Neurais artificiais são essencialmente sistemas compostos de vários níveis de elementos computacionais simples, nós (neurônios), interconectados de acordo com várias arquiteturas ou topologias (Humpert, 1990). A arquitetura destes modelos é especificada pelos nós característicos, topologia da rede e algoritmo de aprendizagem. Um algoritmo de aprendizagem comumente usado é de retropropagação (Rumelhart et al., 1986; Wang & Wu, 1991; Korn, 1990; Jones & Hoskins, 1987; Venkatasubramanian et al., 1990; Parker, 1985; Le Cun, 1985).

Embora várias arquiteturas de redes neurais tenham sido propostas (Lippmann, 1987; Korn, 1990; Wang & Wu, 1991 e Morrison, 1992; Hopfield, 1982; Feldman & Ballard, 1982; Kohonen, 1984; Rumelhart & McClelland, 1986), uma configuração popular de redes neurais por retropropagação é uma rede totalmente “feedforward” (Epping & Nitters, 1990; Venkatasubramanian et al., 1990; Savkovic-Stevanovic, 1991, 1993a,b). Em uma rede “feedforward”, cada neurônio executa uma transformação linear em suas entradas, isto é, cada neurônio coleta e processa a informação, e então a passa adiante. A força da conexão é parametrizada por uma função ponderada que controla o mapeamento da entrada no domínio da saída. A escolha apropriada da função ponderada é feita empiricamente examinando as saídas

preditas com dos dados experimentais ou conhecidos durante o processo de aprendizagem. Os cálculos são feitos sobre a rede inteira, exceto a camada de entrada. O mapeamento de cada unidade é em termos da combinação de todas as entradas (função combinação), seguida pela aplicação de uma função não linear (função de ativação). Para esta, as formas mais comumente usadas são o somatório linear ponderado e a função sigmóide, respectivamente. Outras funções publicadas incluem combinações quadráticas (Kavuri & Venkatasubramanian, 1993; Mah & Chakravarthy, 1992), e uma função de ativação elipsoidal (Kavuri & Venkatasubramanian, 1993).

A estrutura da rede é geralmente determinada pela experiência, com pouco ou nenhum uso de características conhecidas do sistema considerado, e embora tais redes apresentem uma boa performance global, parecem não ter relação ao problema sendo representado e não fornecem entendimento dos padrões de comportamento intrínseco e não são suficientemente seguros quando extrapolando o espaço de aprendizagem dos dados de entrada e saída. O uso de conhecimento prévio pode ser de ajuda na superação destas dificuldades e várias tentativas já têm sido propostas, tais como módulos de diagnose hierárquica (Becraft & Lee, 1993); aglomerados de entradas relacionadas em subredes em uma rede hierárquica (Mavrovouniotis & Chang, 1992); decomposição de tarefa (Raju & Cooney, 1992); e modelos combinados (Wainwright et al., 1994).

Embora redes neurais artificiais não sejam um conceito novo, interesse nesta área de pesquisa tem crescido significativamente nesta última década. A maior razão para este interesse é um resultado da natureza genérica, capacidade de aproximação e adaptabilidade das redes neurais (Cybenko, 1989; Hornik et al., 1989).

Redes Neurais são úteis como módulos em sistemas complexos interagindo muito bem com lógica “fuzzy”, algoritmos genéticos e sistemas especialistas. A combinação cuidadosa de técnicas produz sistemas mais úteis e robustos (Padgett & Roppel, 1992; Crowe & Vassiliadis, 1995).

Redes neurais podem ser empregadas nos mais variados campos da ciência, como em muitos processos de engenharia (vide item a seguir), medicina (ex.: Marshal et al., 1991) economia (ex.: Hoptroff et al., 1991), meteorologia (Schizas et al., 1991), etc.

2.1.1. Aplicação em Engenharia Química

Atualmente, há um forte e crescente interesse dentro da comunidade de engenharia química em computação neural, o que pode ser observado pelo número cada vez maior de publicações descrevendo o uso de redes neurais artificiais em problemas de engenharia química, como detecção de falhas, processamento de sinais, modelagem de processo, e controle (Hoskins & Himmelblau, 1988). Há várias razões para este desenvolvimento. Primeiro, avanços na área de tecnologia de computadores e processamento paralelo têm tornado o uso de redes neurais artificiais mais economicamente possível que no passado. Segundo, desde que redes neurais artificiais são compostas de funções não lineares, elas têm a habilidade de desenvolver bons modelos a partir de dados exemplos e requerem pouco ou nenhum conhecimento prévio da tarefa a ser executada. Terceiro, redes neurais artificiais têm o potencial de resolver certos tipos de problemas complexos que não têm sido satisfatoriamente tratado por métodos mais tradicionais.

Alguns exemplos na literatura de engenharia química de como redes neurais artificiais têm sido aplicadas incluem:

1. retificação de dados: Karjala et al. (1992); Karjala & Himmelblau (1996)
2. detecção e diagnóstico de falhas: Hoskins & Himmelblau (1988); Kramer & Leonard (1990); Ungar et al.(1990); Venkatasubramanian et al. (1990); Watanabe et al. (1989); Naidu et al. (1990); Himmelblau (1991);

Rengaswamy et al. (1995); Becraft & Lee (1993); Kavuri & Venkatasubramanian (1993).

3. modelagem e simulação de processos: Bhat & McAvoy (1990); Su et al. (1992); Lennox et al. (1998); Lambert & Hecht-Nielsen (1991); Montague et al. (1990); Bhat et al. (1989a, b); Rademan et al. (1996); Kaminski et al. (1998); Shaw et al. (1997); Latrille et al. (1994); Zhao et al. (1999); Zhang et al. (1998); Assaf et al. (1996); Cubillos et al. (1996); Leib et al. (1996); Nascimento et al. (1994); Horng et al. (1995); SU et al., 1992
4. controle de processos
 - 4.1. via uso de redes neurais como modelo de processo: Saint-Donat et al. (1991); Montague et al. (1989,1992); Willis et al. (1992); MacMurray & Himmelblau (1995); Bhat & McAvoy (1989, 1990); Pottmann & Seborg (1992); Leonard et al (1992a, b); Blum et al (1992); Ramchandran & Rhinehart (1995); Rohani et al. (1999); Palancar et al. (1998); Ramchandran & Russell (1995)
 - 4.2. para controle direto: Lee & Park (1991);Psichogios & Ungar (1991); Psaltis et al. (1987, 1988); Narendra & Parthasarathy (1990); Ydstie (1990); Miller at al. (1990); Lant et al. (1990); Di Massimo et al. (1990); Gokhale et al.(1995)
 - 4.3. para melhoria de técnicas de controle existentes: Scott et al. (1992); Montague et al. (1989).
5. Otimização de problemas não lineares: Kennedy & Chua (1988); van der Walt et al. (1995); Dong et al. (1996)

6. Identificação: Bhat et al. (1990); Wilson & Zorzetto(1997); Chen & Bruns (1996); Doherty et al.(1997); Pollard et al.(1992); Su et McAvoy (1991), Narendra e Parthasarathy (1990)
7. Termodinâmica: Sharma et al. (1999) Petersen et al. (1994); Spencer et al. (1995)

Além disso, Narendra e Parthasarathy (1990) e Hunt et al. (1992) apresentaram uma revisão das publicações pertinentes ao uso de redes neurais para modelagem, identificação e controle de processos. Bulsari (1994) também apresenta uma revisão do que tem sido feito e do que poderia ser feito aplicando redes neurais em engenharia de processo.

Hussain (1999) fornece uma revisão extensiva das várias aplicações utilizando redes neurais para controle de processos químicos, tanto em simulação e implementação on-line. Esta revisão diz respeito aos três maiores esquemas de controle: controle preditivo, controle baseado modelo inverso e métodos de controle adaptativo, revelando um prospecto valioso do uso de redes neurais em controle de processo. Ele também mostra a rede neural multicamada como a mais popular para tais aplicações, além da falta de sucesso real em aplicações on-line até o momento.

Uma avaliação da aplicabilidade das estruturas de redes neurais como ferramentas de otimização e análise de dados de processo é apresentada por Savkovic'-Stevanovic' (1994).

A maioria dos artigos publicados nesta área considera redes neurais como “caixas pretas” e concentra-se nas aplicações.

2.2. Otimização de Colunas de Destilação

Há numerosos trabalhos e textos disponíveis sobre projetos ótimos de colunas de destilação, mas pouco sobre a operação ótima de uma coluna de destilação existente, que é uma situação mais comum.

Um dos principais problemas na busca de regras gerais acerca de condições de operação ótimas é formular um critério que apresente as condições ótimas ou quase ótimas para um grande número de processos.

As condições ótimas para destilação são discutidas por Shinskey (1977), que sugere minimizar a função custo formada pela somatória do custo da energia consumida para realizar a separação e pela perda do componente chave leve e do componente chave pesado no fundo e no topo da coluna, respectivamente.

Kurt et al. (1978) observam que este critério é adequado para separações onde a qualidade dos produtos não tem uma especificação rigorosa, como por exemplo, em refinarias. Em muitas separações, entretanto, há necessidade de se manter uma especificação mais restrita para ambos os produtos, havendo uma penalidade no preço quando ela não é satisfeita. Para obter um critério genérico para a não satisfação da especificação, um termo relacionado à perda de especificação do produto foi acrescentado ao critério elaborado por Shinskey (1977). Este critério mais geral é usado para definir a operação ótima de uma coluna de destilação e tem como objetivo minimizar as perdas devido a uma recuperação incompleta de produtos, produtos fora de especificação, consumo de energia no condensador e refeedor. Pode ser observado que não há vantagem em se produzir produtos mais puros que o valor especificado. Segundo Shinskey (1977) esta é a situação mais comumente encontrada.

Um dos pontos cruciais na modelagem de processos é a sua complexidade. Em cada aplicação há um grau de detalhamento adequado na modelagem e por consequência na precisão obtida. No trabalho desenvolvido por Kurt et al. (1978)

foram feitas as seguintes simplificações: separação binária, cálculos realizados no estado estacionário de acordo com McCabe-Thiele, eficiência independente da hidráulica dos pratos e pressão constante.

Kurt et al. (1978) definem “regras” que indicam se a coluna de destilação está sendo operada no ponto ótimo.

Edgar & Himmemblau (1989) classificam os problemas de otimização de colunas de destilação em estado estacionário em três categorias, em ordem crescente de complexidade:

- Determinação das condições operacionais ótimas de uma coluna de destilação existente a uma dada especificação, conhecidas as características da alimentação.
- Otimização do projeto de uma coluna em que, além de determinar as condições operacionais ótimas, é também necessário calcular o tamanho da coluna.
- O outro problema é, além de calcular o número de estágios, determinar a localização ótima da alimentação e das retiradas laterais.

Um exemplo de otimização de uma coluna de destilação convencional existente é dado em Edgar & Himmemblau (1989).

Srygley & Holland (1965) propuseram um método de otimização de projeto, para determinar o número mínimo de estágios para colunas de destilação convencionais (uma alimentação, produto de fundo e de topo) e complexas (várias alimentações e retiradas laterais), obedecendo a um conjunto de especificações de pureza do produto.

Martin, Latour e Richard (1981) apresentam a dualidade entre custos de energia e grau de separação. Se o custo de energia é menor que um dado valor limite, o procedimento ótimo de operação é maximizar o refluxo.

Segundo Fauth e Shinsky (1975), para a grande maioria das colunas de destilação, o controle da qualidade do produto exatamente no limite da especificação resultará em mínimo custo operacional. Este ótimo local pode ou não ser consistente com a otimização do sistema. Os autores procuram mostrar as vantagens econômicas da aplicação do controle regulatório conjuntamente com a otimização local de uma coluna depropanizadora. Aqui, novamente, o equacionamento do custo é função do valor dos produtos recuperados e da energia consumida na separação.

Otimização “on-line” utiliza todo o potencial da otimização estática já que a condição operacional ótima não permanece constante com o tempo, alterando-se à medida que o preço dos produtos e do combustível muda.

2.3. Azeotropia

Azeótropo, uma palavra de origem grega que é traduzida como “ebulição constante”, significa que, para uma dada mistura, o vapor emitido tem a mesma composição do líquido (Widagdo & Seider, 1996). Um conhecimento de azeótropos é desejável por duas razões: primeiro, eles geralmente ocorrem em fracionamento e torna uma dada separação impossível por retificação simples; segundo, eles podem ser utilizados para separar misturas não comumente separadas por retificação direta. A formação de misturas com ebulição constante é uma função da diferença nas temperaturas de ebulição dos componentes puros e da não idealidade da solução resultante. A não idealidade de uma mistura depende fortemente das forças intermoleculares dos componentes tais como, forças de dispersão, interações dipolo - dipolo, interações dipolo – dipolo induzido, pressão interna e pontes de hidrogênio, dos quais o último é o efeito mais importante na formação de um azeótropo (Ewell et al., 1944; Kim & Simmrock, 1997).

Até o presente, muitas tentativas têm sido feitas para prever e correlacionar o fenômeno de azeotropia. Um grande número de dados experimentais tem sido obtido

e compilado em bancos de dados. Um dos mais importantes foi adicionado Banco de Dados de Dortmund, o qual contem cerca de 40000 informações sobre o comportamento azeotrópico de misturas binárias e ternárias. Todos os dados (35000) armazenados antes de 1993 foram publicados após cuidadosa avaliação em uma forma unificada como uma compilação de dados (Gmehling et al., 1994). O procedimento para checar os dados é descrito em detalhes nas seguintes referências: Gmehling et al. (1994, 1995). Outras compilações de dados azeotrópicos são Lecat (1949) e Horsley (1973).

Um banco de dados computadorizados tem sido estabelecido na Universidade de Oldenburg para suplementar a informação de mistura armazenada no Banco de Dados de Dortmund (Gmehling et al., 1995; Gmehling & Bölts, 1996).

Embora exista considerável interesse na área de predição de equilíbrio de fases para misturas, relativamente poucos métodos para a predição de azeótropo têm sido publicados. São várias as explicações: dificuldade em determinar a pressão, temperatura, e composição de *todos* os azeótropos, devido à forma altamente não linear das equações que constituem os modelos termodinâmicos (Harding et al., 1997), falta de um modelo termodinâmico completamente satisfatório para o estado líquido e falta de conhecimento suficiente das forças intermoleculares em uma mistura (Kim & Simmrock, 1997).

A maioria dos trabalhos publicados na literatura tem sido limitada a cálculo de azeótropos homogêneos. Aristovich & Stepanova (1970) calcularam azeótropos homogêneos ternários usando o modelo de Wilson sob condições isotérmicas. Teja & Rowlinson (1973) calcularam azeótropos homogêneos de misturas binárias usando equações de estado como modelo termodinâmico. Sua proposta foi fixar temperatura e variar composição e volume até que as condições de equilíbrio termodinâmico fossem satisfeitas. Wang & Whiting (1986) também usaram uma equação de estado para calcular azeótropos homogêneos para misturas binárias. Trotsenko (1997) apresentou um método para predição e cálculo do comportamento azeotrópico sem executar cálculos de equilíbrio líquido - vapor, empregando em seu lugar a equação

de Maxwell que pode exibir um único extremo no ponto azeotrópico. O uso do método é demonstrado usando equações de estado cúbicas.

Devido à dificuldade em fornecer garantias quanto à boa predição da localização do azeótropo empregando modelos de comportamento de fase, tem sido considerável o interesse no desenvolvimento de técnicas mais confiáveis para o cálculo de azeótropos homogêneos. Por exemplo, Fidkowski et al. (1993) e Westerberg & Wahnschafft (1996) apresentaram métodos contínuos de homotopia com esta finalidade, e Okasinski & Doherty (1977) usaram um método similar para encontrar azeótropos reativos. Embora estes métodos sejam muito seguros, uma desvantagem é que eles não podem garantir completamente que todos os azeótropos sejam determinados.

Maranas et al. (1996) e Harding et al. (1997) usaram um poderoso procedimento de otimização global: ele é baseado na “estrutura de árvore” com funções estimadoras convexas que são continuamente refinadas à medida que o domínio no qual o azeótropo pode ocorrer é reduzido. Esta técnica fornece uma garantia matemática que todos os azeótropos são incluídos, embora esta garantia possa ser perdida devido a erros computacionais de arredondamento. Harding et al. (1997) desenvolveram funções estimadoras convexas para vários modelos termodinâmicos específicos.

Maier et al. (1998) descreveram um método para seguramente incluir todos os azeótropos homogêneos de misturas multicomponentes, e verificar quando não existem. A técnica é baseada na análise do intervalo, em particular o uso de um algoritmo Newton/bisseção generalizado. O método é matematicamente e computacionalmente seguro para determinar quaisquer e todos os azeótropos homogêneos, tratando automaticamente com erros de arredondamento, não requer ponto inicial de partida e pode ser aplicado em conexão com qualquer modelo termodinâmico.

Chapman & Goodwin (1993) apresentaram um método de busca para localizar azeótropos homogêneos e heterogêneos que usa um algoritmo de Levenberg-Marquardt para encontrar azeótropos homogêneos e então verifica a estabilidade de cada solução com o critério do plano tangente descrito por Michelsen (1981). Uma solução instável é, então, usada como ponto de partida para uma nova busca para um azeótropo heterogêneo. Este método é limitado pelo fato que azeótropos heterogêneos não necessariamente correspondem a uma solução azeotrópica homogênea instável e vice-versa.

Uma excelente revisão de destilação não ideal, incluindo uma discussão sobre determinação de azeótropos foi publicada por Widagdo & Seider (1996).

Kim & Simmrock (1997), criou um sistema especialista, chamado AZEOPERT, com base em regras hierárquicas para predizer a formação de azeótropos binários entre componente em uma mistura.

3. REDES NEURAIS

3.1. Introdução

Rede Neural Artificial é frequentemente identificada como uma subespecialidade da Inteligência Artificial. Foi muito utilizada para problemas de classificação e reconhecimento de padrões, como parte da teoria conexionista dos processos mentais e também como uma categoria de modelos em ciência da cognição.

As redes neurais são funções que estimam relações de entrada – saída de um dado sistema. São, portanto, mais uma ferramenta a ser utilizada em modelagem de processos. Sua característica mais interessante é que não dependem de um modelo matemático que relacione as entradas e as saídas de um processo. Elas “aprendem” essa relação a partir de “treinamento” semelhante ao aprendizado de um cérebro humano. Em função dessa característica, a rede neural é aplicada em casos que apresentam forte não-linearidade, o que dificulta bastante a obtenção de modelos fenomenológicos.

Embora existam muitos tipos de redes neurais (RN), duas classes são geralmente empregadas em engenharia química:

- ✓ rede de múltiplas camadas, cuja informação segue uma única direção, “multilayer feedforward network” (Fig. 3.1)
- ✓ rede recursiva, “recurrent network” (Fig. 3.2), cujas informações de saída retornam à entrada da RN.

As RNs de múltiplas camadas possuem características de modelar bem processos estacionários enquanto que as RN recorrentes são melhores para modelar processos dinâmicos.

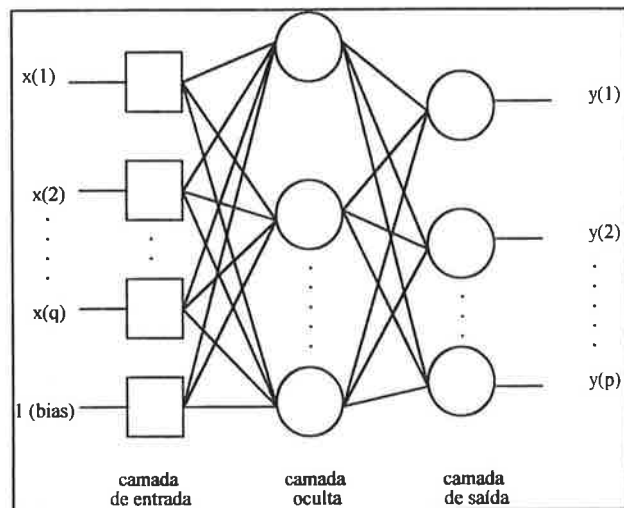


Fig. 3.1 - Rede de Múltiplas Camadas - “Multilayer Feedforward Network”

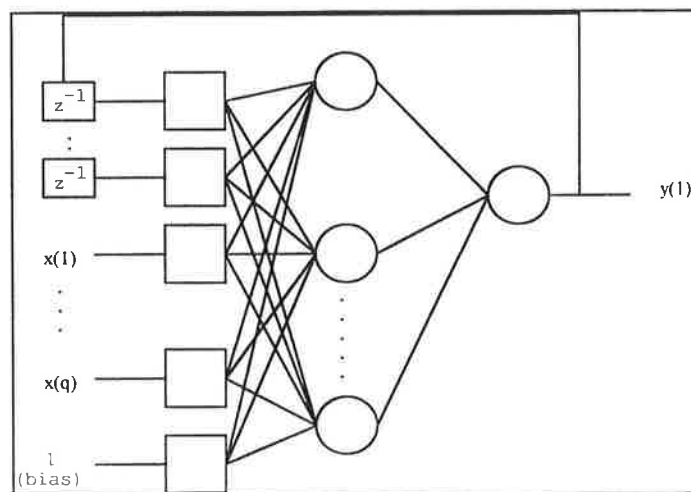


Fig. 3.2- Rede Recursiva - “Recurrent Network”

A rede neural comumente empregada na modelagem e simulação de processos químicos é a rede “feedforward”, RNFF. Os nós nas diferentes camadas da rede representam elementos de processamento do tipo neurônios. O número de neurônios nas camadas de entrada e saída depende dos respectivos números de dados de entrada e saída considerados. Por outro lado, neurônios na camada interna podem variar, dependendo da especificação do usuário, e sua estrutura define a topologia da rede “feedforward”. Os neurônios de uma camada são conectados a todos os neurônios da camada seguinte. Cada interconexão tem associado um peso que modifica a força do

sinal fluindo através do caminho. Assim, a entrada de cada neurônio é uma soma ponderada das saídas dos neurônios da camada anterior. Em outras palavras, cada informação que sai de um neurônio de uma camada i é ponderada por um dado peso W_{ij} e enviada a todos os neurônios da camada seguinte j . A saída de cada nó é obtida passando a soma ponderada através de um operador linear chamado função de ativação. Neste trabalho foi usada a função sigmóide, que pode ser considerada como um aproximador universal e é capaz de ajustar funções multivariáveis côncavas e convexas (Cybenko, 1989). Esta função confere a rede habilidade de representar relações não-lineares. A convergência da rede neural está ligada à melhor escolha possível do conjunto de pesos W .

A rede escolhida possui três camadas: a camada de entrada de dados, a camada oculta (NH), “hidden layer”, e a camada de saída. Os neurônios na camada de entrada não efetuam nenhuma função de processamento; eles apenas fornecem o meio pelo qual os dados são introduzidos na rede. A seleção de uma estrutura ótima de rede é crucial para o desempenho da rede. Por exemplo, uma rede com poucos neurônios na camada oculta pode ser capaz de representar o sistema satisfatoriamente e uma rede com muitos neurônios na camada oculta pode apresentar um sobre ajuste no conjunto de treinamento.

Uma vez que a topologia da rede é especificada, um conjunto de dados de entrada e saída é usado para treinar a rede, isto é, para determinar valores apropriados para os pesos associados com cada interconexão. Os dados são propagados em uma única direção através da rede para produzir uma saída que é comparada com o correspondente conjunto de dados experimentais de saída apresentados à rede, gerando, desta forma, um erro. Este erro é minimizado variando os pesos, e pode envolver várias passagens do conjunto de dados de treinamento através da rede. Para uma dada topologia, a magnitude dos pesos define as características e as propriedades estruturais do modelo. Assim, uma rede neural tem a capacidade de representar sistemas complexos cujas propriedades estruturais sejam desconhecidas (Willis et al., 1991).

Há um grande número de parâmetros (pesos) envolvidos na caracterização das redes neurais. Estes parâmetros refletem a alta conectividade entre os neurônios. A maneira típica de definir a estrutura de uma rede consiste em treinar várias redes diferentes com uma parte dos dados, chamado de conjunto de treinamento, e seu desempenho é avaliado, em seguida, com a parte remanescente dos dados, chamado conjunto de teste. Ambos os conjuntos devem conter dados suficientes para assegurar satisfatoriamente o treinamento dos pesos da rede e a avaliação precisa de seu desempenho (Schenker and Agarwal, 1996). O conjunto de treinamento usado para ajustar os dados deve, preferencialmente, conter os extremos de todas as variáveis de entrada e saída, e o conjunto de teste usado para verificar a capacidade de predição da rede deve estar contido no domínio do conjunto de treinamento.

O objetivo de treinar uma rede é encontrar um conjunto ótimo de pesos. Quando o número de pesos é maior que o número de pontos de dados disponíveis, o erro no ajuste dos dados não treinados inicialmente diminui, mas em seguida aumenta à medida que a rede torna-se sobre treinada. Ao contrário, quando o número de pesos é muito menor que o número de dados, o sobre ajuste não é crucial.

Para desenvolver uma rede neural pode-se contar com várias ferramentas, entre elas:

- Programas comerciais
- Simuladores comerciais
- Auxiliares matemáticos comerciais (MATLAB)
- Programas gerados a partir do equacionamento disponível na literatura, por exemplo, o programa NEURO14 (Nascimento, 1992), utilizado neste trabalho.
- Placas de microcomputadores com “chips” neurais.

3.2. Metodologia genérica e seqüencial de treinamento de RN

A modelagem por redes neurais consiste basicamente na determinação dos pesos associados a cada entrada de dados da rede, que minimizam o erro representado pelo quadrado da diferença entre os valores experimentais e os calculados pela RN, conforme a equação:

$$\min(E) = \min \left[\sum_{m=1}^r \sum_{k=1}^p (y_k^{(m)} - O_k^{(m)})^2 \right] \quad (3.1)$$

onde: E – resíduo quadrático

p – número de neurônios da camada de saída

r – número de pontos do conjunto de treinamento

y_k – valor experimental

O_k – valor calculado pela RN na camada de saída

O termo utilizado para expressar a seqüência de cálculos realizados para a determinação desses pesos é “treinamento da rede” e o conjunto de variáveis de entradas e saídas da rede, utilizado nesta fase, é denominado “dados de treinamento”. Para testar a consistência do ajuste, utiliza-se um segundo grupo de variáveis, os “dados de teste”. A validação da rede é feita através da simulação de um terceiro conjunto de variáveis, não empregado no treinamento ou no teste, evitando-se, desta forma, problemas de sobreajuste.

As principais etapas para o desenvolvimento de um modelo de rede neural (RN) e ajustes realizados com este modelo podem ser descritas de forma seqüencial como:

- Análise do processo
- Construção de base de dados

- Configuração de uma estrutura da RN
- Treinamento e aprendizado da RN
- Testes de validação da RN

a) Análise do problema a ser modelado através de RN

Todas as variáveis que caracterizam o processo devem ser consideradas, inclusive as geométricas, pois o procedimento de treinamento assemelha-se ao processo de aprendizado do cérebro humano, na falta de informações importantes a RN vai determinar relações não lineares também de forma incompleta tornando sua capacidade preditiva muito pobre. Por outro lado, deve-se tomar cuidado para não se utilizar informações redundantes.

Nesta etapa, também é feito o levantamento das restrições operacionais, ambientais, de segurança, de especificação de produto, bem como dos limites máximos e mínimos das variáveis envolvidas, e sensibilidades das mesmas, de modo a auxiliar na obtenção da superfície de resposta do processo, visando à otimização do mesmo.

b) Coleta de dados para o treinamento da RN

Não só a qualidade como também a quantidade de dados relativos ao processo é de fundamental importância. Valores com confiabilidade discutível não só atrapalham o treinamento da RN como também poderá levar a predições futuras de baixa qualidade, senão errôneas.

c) Verificação do domínio das variáveis da RN

Redes Neurais apresentam ótima característica de interpolação de resultados, logo é capaz de simular casos contidos dentro do domínio do conjunto de dados empregados no conjunto de aprendizagem. No entanto, dados fora deste domínio (extrapolação) não são bem representados e devem ser evitados ou acompanhados com cuidado. Desta maneira, torna-se um procedimento sistemático e importante que o domínio do conjunto de aprendizagem contenha o domínio do conjunto de teste e/ou simulação.

d) Normalização dos dados

Em geral, a faixa de normalização dos dados encontra-se entre [0, 1]. Aconselha-se, no entanto, a utilização de um intervalo menor. Para os casos estudados neste trabalho, utilizou-se o intervalo de normalização entre [0,1, 0,9] de modo a manter os valores dos pesos w_{ij} mais equilibrados e não muito altos.

e) Treinamento da RN

Como as entradas e saídas da RN já foram definidas anteriormente, sabe-se, de antemão, o número de neurônios nas camadas de entrada e saída da RN. Após esta definição e de posse de um software específico, arbitra-se o número de neurônios na camada oculta, o número de iterações para o treinamento, o valor do resíduo e os conjuntos de dados para treinamento e teste, procedendo-se, em seguida, ao treinamento propriamente dito. De posse de várias configurações de RN, escolhe-se aquela que apresentou menor valor de erro residual para o conjunto do teste.

Na escolha dos pesos correspondentes ao erro mínimo, podem ocorrer, basicamente, duas situações:

- *Caso 1:* o erro descrito pela equação acima, obtido a partir dos dados de treinamento e teste diminui em função do número de apresentações. Neste caso, o número de apresentações correspondente ao erro mínimo é o mesmo para os dois conjuntos de dados e está, portanto definido o conjunto de pesos da rede neural.
- *Caso 2:* o erro obtido a partir dos dados de teste pode apresentar um mínimo. Nesse caso, o conjunto de pesos escolhidos para a rede é o correspondente ao número de apresentações dos dados de teste que apresentam o menor erro.

A fase de aprendizado ou treinamento é basicamente um problema de otimização e freqüentemente exige a apresentação da base de dados à rede por milhares de vezes. Dessa forma, depara-se com as mesmas dificuldades associadas à otimização convencional de processos, tais como: convergência, existência de mínimos locais, esforço computacional, entre outras. Não existe, até hoje, um algoritmo comprovadamente eficiente para assegurar um treinamento ótimo.

A metodologia para a fase de treinamento de redes neurais foi basicamente consolidada por RUMELHART e MCCLELLAND (1986), conhecida como método da retropropagação (“backpropagation”). Tal metodologia foi obtida com base no método de otimização do gradiente descendente e tem sido utilizada na grande maioria dos trabalhos publicados na área de redes neurais, aplicadas a processos químicos.

f) Validação da RN treinada

Após a seleção da RN que apresenta o melhor resultado, isto é, que melhor representa as variáveis de saída, passa-se aos testes de consistência. O primeiro teste a ser realizado é em relação à superfície de resposta de uma variável da camada de saída em função das variáveis de entrada da RN. Um critério

sinalizando um bom ajuste está na obtenção de superfícies regulares. Qualquer variação drástica pode significar erro de sobreajuste (Pollard et al., 1992). O segundo teste de consistência diz respeito à distribuição da função distância (definida como a diferença entre os valores experimentais e calculados, $(y_{exp} - y_{calc})_{RN}$ ou análise da dispersão dos valores calculados das variáveis de saída em comparação aos valores medidos. Estas informações fornecem a distribuição da diferença do erro absoluto e avaliam a boa qualidade do ajuste e a distribuição normal demonstra o equilíbrio de frequência dos valores encontrados em torno da média zero. Uma distribuição normal mostra a boa qualidade do ajuste realizado.

g) Utilização da RN

Uma vez treinada, a RN pode ser simulada com grande rapidez. Milhares de simulações podem ser processadas em segundos, o que possibilita o emprego do modelo de redes neurais em procedimentos de modelagem, simulação, otimização, controle de processos e controle em linha, análise de restrições do processo, análise de sensibilidade das variáveis, analisadores virtuais, predição de propriedades termodinâmicas, entre outras aplicações.

Neste trabalho, para treinamento da rede foi utilizado o programa NEURO14, desenvolvido no Laboratório de Simulação e Controle de Processos do Departamento de Engenharia Química/Escola Politécnica da USP (Nascimento, 1992).

A rede treinada é do tipo “feedforward”, com algoritmo de retropropagação (“backpropagation”) para convergência.

3.3. Aplicações em Engenharia Química

No contexto de aplicações em engenharia química, redes neurais apresentam particular interesse como modelos preditivos (Nascimento et al., 1998, 1999 e 2000) e reconhecimento de padrões (Nascimento et al., 1997). Além disso, redes neurais têm sido largamente usadas como ferramentas em áreas tais como análise, simulação e controle de processos devido principalmente a suas vantagens inerentes tais como adaptabilidade, não-linearidade, e relativa simplicidade de uso. Redes neurais possuem a habilidade de aprender o comportamento do processo sem realmente especificar as leis físicas e químicas que regem o sistema. Em adição, modelos de redes neurais podem simultaneamente empregar variáveis contínuas e discretas (Migliavacca et al., 1999). O sucesso em obter uma rede robusta e confiável depende fortemente da escolha de variáveis apropriadas, bem como da qualidade do conjunto de dados disponíveis e do domínio usado para fins de treinamento.

Uma característica dos processos químicos é que a maioria deles apresenta forte não linearidade, o que dificulta a obtenção de modelos fenomenológicos, a ponto de, em alguns casos, conseguir-se apenas um modelo extremamente simplificado, válido em condições muito específicas. É exatamente nessa área que as redes neurais surgem como uma técnica interessante a ser empregada, devido à sua característica de “aprender” o que ocorre no processo. Pode-se contornar, assim, em muitos casos, as dificuldades de obtenção de modelos fenomenológicos representativos do processo. Entretanto, o sucesso da modelagem via rede neural depende fortemente do conhecimento das variáveis significativas do processo, além de ser necessária uma base de dados de boa qualidade, ou seja, que contenha todas as informações importantes do processo no domínio desejado.

3.4. Estrutura Matemática da Rede e o Algoritmo de Aprendizado

Como mencionado previamente, a rede neural comumente empregada na modelagem e simulação de processos químicos é a rede “feedforward”, RNFF. Nesta

rede os neurônios de uma camada são conectados a todos os neurônios da camada seguinte. Cada informação X que sai de um neurônio de uma camada i é ponderada por um dado peso W_{ij} e enviada a todos os neurônios da camada seguinte j . A convergência da rede neural está ligada à melhor escolha possível do conjunto de pesos W . A rede escolhida possui três camadas: a camada de entrada de dados, a camada oculta (NH), “hidden layer”, e a camada de saída. A camada de entrada não efetua nenhum processamento.

A fim de diminuir as dificuldades de convergência, todos os dados do conjunto de aprendizagem são normalizados. Além dos dados de entrada, alimenta-se também um valor de referência “bias”, à rede. A figura 3.3 apresenta um neurônio de uma camada j da estrutura da rede.

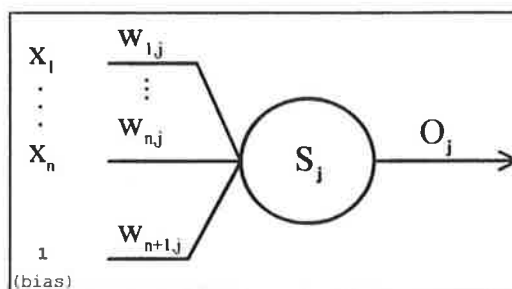


Fig. 3.3. - Representação de um neurônio da camada j

Todas as variáveis de entrada X_i são ponderadas com um valor W_{ij} e somadas, conforme a eq. 3.2:

$$S_j = \sum_{i=1}^n w_{i,j} X_i + w_{n+1,j} \quad (3.2)$$

A saída do neurônio, O_j , é calculada pela expressão:

$$O_j = f(S_j) \quad (3.3)$$

A relação entre a entrada e a saída dos neurônios é dada por uma função do tipo sigmóide:

$$f(\alpha) = \frac{1}{1 + e^{-\alpha}} \quad (3.4)$$

HUNT et al. (1992) apresentam outras formas desta relação $f(\alpha)$ utilizadas em modelagem e controle de processos químicos, bem como um conjunto de recomendações quanto à utilização de um ou outro tipo de rede neural.

Dependendo da função de ativação escolhida, o intervalo de normalização das funções pode mudar. O problema em se diminuir muito o intervalo de normalização dos dados é que a RN perde sua principal característica que é a capacidade de mapeamento de sistemas não lineares, comportando-se como uma regressão multilinear.

O treinamento de uma rede neural consiste de uma adaptação de pesos, W_{ij} , aplicados sobre as informações entre as camadas de neurônios, de forma a minimizar o erro quadrático, expresso na forma, por exemplo, da função E, mostrada na eq. 3.5. Esta função representa o valor acumulado para todos os dados experimentais, de 1 a γ , das somas dos erros quadráticos entre os valores experimentais e os calculados pela rede neural, para todas as saídas da rede, de 1 a p :

$$\min(E) = \min \left[\sum_{m=1}^{\gamma} \sum_{k=1}^p (Y_{k,calc} - Y_{k,exp})^2 \right] \quad (3.5)$$

Outros critérios foram recentemente propostos na literatura, como os de POLLARD et al., 1992, nos quais a função objetivo seria menos sensível a dados com erros grosseiros.

Para se obter o conjunto de pesos, W_{ij} , é preciso ter, em primeiro lugar, um conjunto de dados formado por subconjuntos de pares, entrada e saída, do processo, que constituem o conjunto de aprendizado. Este conjunto deve ser montado de tal forma que se contemplem todas as informações relevantes do processo. Tendo esse conjunto de dados, utiliza-se um algoritmo de otimização, a fim de se obter o conjunto de pesos que satisfaça um critério dado, por exemplo, a função apresentada pela eq. 3.5.

O algoritmo mais utilizado para determinar o conjunto de pesos, W , é o de retropropagação, que é uma derivação do gradiente descendente (RUMMELHART e MCCLELLAND, 1986) e tem sido utilizado na grande maioria dos trabalhos publicados na área de redes neurais, aplicados a processos químicos. Este é obtido a partir da minimização do resíduo quadrático, E , com respeito aos pesos W_{ij} . Desta forma, o gradiente $\partial E / \partial W_{ij}$ pode ser calculado e, conseqüentemente, a direção a ser seguida no processo de minimização do resíduo. Em cada passo da otimização, m , a atualização dos pesos pode ser dada pela eq. 3.6:

$$W_{ij}^{(m+1)} = W_{ij}^{(m)} + \Delta W_{ij}^{(m)} \quad (3.6)$$

Esta atualização pode ser feita, alternativamente, utilizando a técnica do “momento”, baseada na inclusão do termo, α , proporcional à variação dos pesos, de modo que $0 < \alpha < 1$ na eq. 3.7. O método dos momentos leva em consideração o caminho do treinamento passado.

$$W_{ij}^{(m+1)} = (1 - \alpha)W_{ij}^{(m)} + \alpha\Delta W_{ij}^{(m)} \quad (3.7)$$

A diferença entre o método do gradiente descendente e o “backpropagation” é que este último usa apenas um subconjunto de dados entrada/saída de cada vez, enquanto o método do gradiente tradicional emprega todos os dados do conjunto de aprendizado. Para a camada de saída, a variação dos pesos referentes ao neurônio de saída k , incidentes sobre a saída do neurônio j , da camada oculta, a cada interação m , é dada por:

$$\Delta W_{ij}^{(m)} = \eta * f'(S_k)(Y_k^{(m)} - O_k^{(m)})O_j^{(m)} \quad (3.8)$$

Para a camada oculta, a variação dos pesos referentes ao neurônio oculto j incidente sobre a entrada X_i , a cada interação m , é dada por:

$$\Delta W_{ij}^{(m)} = \eta * f'(S_j) \left\{ \left[\sum_{k=1}^p f'(S_k)(Y_k^{(m)} - O_k^{(m)})W_{ik}^{(m-1)} \right] X_i^{(m)} \right\} \quad (3.9)$$

Nas duas equações acima, f' é a derivada da função sigmoideal, dada por:

$$f'(z) = [e^{-z} / (1 + e^{-z})]^2 \quad (3.10)$$

e η é um fator de amortecimento ou aceleração, que pode ser utilizado para evitar passos excessivamente grandes no processo de minimização, que podem causar divergência da função objetivo. Existem trabalhos que sugerem que este fator tenha

valores menores que 0,2 e outros, valores maiores, inclusive maiores que a unidade. Existem trabalhos que tentam acelerar o treinamento quando o resíduo quadrático tem tendência a diminuir e a desacelerar o treinamento quando o resíduo quadrático tem tendência a aumentar (Nascimento et al., 1993). Em RNs também é possível melhorar significativamente a qualidade do treinamento alterando-se a forma com que o fator de amortecimento varie ao longo do treinamento.

No algoritmo de retropropagação, em cada iteração, são calculadas inicialmente as saídas da camada oculta, O_j e da camada de saída O_k , com os pesos estimados na iteração anterior pelas eqs. 3.2 a 3.4 para os n neurônios da camada oculta e p neurônios da camada de saída. Em seguida, as saídas da rede são utilizadas para o cálculo dos novos valores dos pesos, utilizando-se as eqs. 3.8 a 3.10 para as variações e a eq. 3.6 ou 3.7 para os novos valores. Isto é feito para cada apresentação dos pontos experimentais. Assim, inicialmente os erros são propagados para frente na rede e utilizados para estimar os novos valores dos pesos nas camadas anteriores.

Uma vez determinado o conjunto de pesos W que minimiza o resíduo E , para simular o processo basta multiplicar um conjunto de dados de entrada qualquer pelos respectivos pesos, respeitando-se a estrutura da rede, obtendo-se o conjunto de resultados após a camada de saída da rede.

4. OTIMIZAÇÃO QUALITATIVA/QUANTITATIVA DE PROCESSOS

INDUSTRIAIS

4.1. Introdução

Análise ou simulação de processo é a ferramenta que o engenheiro químico usa para interpretar fluxogramas, para localizar o mau funcionamento e para prever a performance de processos. O coração da análise é o modelo matemático, um conjunto de equações que relaciona as variáveis do processo, tais como temperatura, pressão, vazão e composição de uma corrente, área superficial, características das válvulas, configurações geométricas, entre outros.

Os simuladores de processo em estado estacionário calculam variáveis desconhecidas, uma vez dados os valores de certas grandezas. Isto é, dado um fluxograma de processo detalhado para o projeto caso base, ou um fluxograma de processo depois de uma etapa de integração em uma síntese de processo, ou mesmo um fluxograma incompleto depois de uma das etapas preliminares, é geralmente possível usar um simulador para encontrar valores desconhecidos para temperaturas, pressões e vazões em estado estacionário. Para um processo existente, análises usando um simulador de processo são realizadas rotineiramente para estudar potenciais mudanças nas condições de operação ou a possibilidade de um redimensionamento para melhorar sua lucratividade.

Há vários níveis de análise. Em ordem de complexidade crescente, eles envolvem: balanços materiais, balanços de material e energia, dimensionamento de equipamentos, e análise econômica. Equações adicionais são implementadas em cada nível. Novas variáveis são introduzidas, e os algoritmos tornam-se mais complicados.

Felizmente, a maioria dos processos químicos envolve equipamentos convencionais de processo: trocadores de calor, bombas, colunas de destilação, absorvedores, entre outros. Para estas unidades de processo, as equações não diferem

entre os processos químicos. As propriedades físicas e termodinâmicas e as constantes cinéticas diferem, mas não as equações que regem os equipamentos. Assim, é possível preparar um ou mais algoritmos para cada unidade de processo de modo a resolver os balanços de material e energia e para calcular o tamanho dos equipamentos e custos. Uma biblioteca de sub-rotinas ou modelos que automatizam tais algoritmos faz parte do coração de simuladores de processos.

Os modelos de estado estacionário em simuladores não estão aptos para resolver equações dependentes do tempo. Eles simulam operações em estado estacionário de unidades de processo (sub-rotinas de operação) e estimam as dimensões e os custos das unidades de processo (sub-rotinas de custos). Dois outros tipos de sub-rotinas são usados para otimização e/ou cálculos de convergência de ciclos (sub-rotinas de convergência) e para alterar os parâmetros dos equipamentos (sub-rotinas de controle). Para levar em conta a dinâmica do processo (dependência com o tempo), faz-se uso de simuladores dinâmicos.

Quando se utilizam simuladores de processos, é importante reconhecer que, com algumas suposições, a maioria das correntes é constituída por espécies químicas que se distribuem em uma ou mais fases que são assumidas estar em equilíbrio. Então, é importante, quando empregar simuladores de processo, compreender como eles aplicam a teoria do equilíbrio de fases na modelagem de correntes, e unidades de processo que representam operações de equilíbrio (vasos, colunas, escoamento bifásico, trocadores, entre outros).

Atualmente, simuladores comerciais modelam uma ampla faixa de processos petroquímicos e desde tradicionais processos químicos à química fina e especialidades, metais e minerais, processos alimentícios e de celulose, gerenciamento ambiental, processos farmacêuticos e biológicos.

Como dito anteriormente, simuladores comerciais são ferramentas importantes e largamente utilizadas na operação de plantas existentes e em novos projetos. Além disso, já é possível utilizar estas ferramentas como algoritmos automáticos de

otimização a fim de encontrar o melhor projeto ou condições operacionais, isto é um simulador pode calcular um ponto ótimo que satisfaça uma dada função objetivo pré-definida com todas as suas restrições. Assim, métodos formais de otimização podem ser utilizados para otimizar uma super estrutura de unidades de processo com correntes que podem ser consideradas ou não, usando variáveis binárias. A formulação mista inteira do problema de otimização, em princípio, permite ao otimizador selecionar simultaneamente o melhor fluxograma e otimizá-lo em relação a suas variáveis contínuas, tais como níveis de pressão, razões de refluxo, tempos de residência, e frações de separação. Na prática, entretanto, a maioria dos problemas de projeto não é resolvida usando super estruturas e algoritmos de otimização mista inteira. Preferivelmente, regras heurísticas em conjunto com métodos de simulação são utilizadas para construir e analisar árvores de síntese. Embora super estruturas, tais como redes de trocadores de calor, possam ser otimizados convenientemente utilizando métodos do tipo misto inteiro, é impraticável tentar a otimização de um processo complexo desta maneira (Seider et al., 1999). Por exemplo, uma coluna de destilação modelada considerando estágios de equilíbrio geralmente tem centenas de equações MESH a serem resolvidas. No entanto, na implementação da maioria dos simuladores, estas equações são resolvidas para cada unidade do processo, parâmetros de equipamentos dados e variáveis de correntes (geralmente para correntes de alimentação), usando sub-rotinas de uma biblioteca de programas. Dessa forma, quando se utilizam estes simuladores, as equações (ou restrições de igualdade) para as unidades de processo são mostradas explicitamente em um programa não linear. Dados valores para as variáveis de projeto, os simuladores chamam estas sub-rotinas para resolver as equações apropriadas e obter as incógnitas necessárias para executar a otimização. Este método consome tempo uma vez que, geralmente, requer um alto número de iterações. Além disso, no projeto e otimização de processo, cerca de 80% do tempo computacional total pode ser gasto na avaliação de propriedades termodinâmicas (Seider et al., 1999).

Ainda que otimização sem utilização de modelo por experimentação direta seja possível (usando técnicas de planejamento de experimentos), os custos e esforços envolvidos podem ser significativamente altos, especialmente quando o número de

variáveis é grande. Então, um importante pré-requisito é a disponibilidade de um modelo matemático válido e representativo. Quando disponíveis, modelos do primeiro princípio podem ser uma escolha natural. Entretanto, na maioria dos problemas de engenharia química, tais modelos são não lineares e computacionalmente demorados. Com o objetivo de otimização, o que envolve muitas simulações e particularmente aplicações em linha, é muito importante que o modelo possa ser simulado rapidamente e ao mesmo tempo ser capaz de descrever muito bem o sistema. Neste sentido, alguns trabalhos têm sido propostos com o uso de redes neurais artificiais em substituição aos modelos de primeiro princípio. Redes Neurais têm sido aclamadas como um aproximador não linear universal, e sua aplicação em casos reais de produção tem crescido rapidamente.

Casos que apresentam complexidade do sistema industrial ou falta de conhecimento da tecnologia fundamental característica do processo considerado dificilmente podem ser modelados aplicando procedimentos baseados no primeiro princípio. Uma característica dos processos químicos é que a maioria deles apresenta forte não linearidade, o que dificulta a obtenção de modelos fenomenológicos, a ponto de, em alguns casos, conseguir-se apenas um modelo extremamente simplificado e válido somente em condições muito específicas. Para tais casos, modelos de redes neurais podem ser empregados, devido a sua característica de “aprender” o que ocorre no processo com base em dados históricos e/ou experimentos planejados da planta. Entretanto, o sucesso da modelagem via redes neurais depende fortemente do conhecimento das variáveis significativas do processo, da geometria do sistema, além de ser necessária uma base de dados de boa qualidade que contenha todas as informações importantes do processo no domínio desejado. A qualidade da informação é fundamental. A vantagem dos modelos de redes neurais sobre os modelos empíricos irá depender fortemente do grau de não linearidade do processo. Modelos de redes neurais são funções que estimam relações de entrada e saída de um sistema, constituindo-se, portanto, mais uma ferramenta a ser utilizada na modelagem de processos. Sua característica mais interessante é a não dependência de um modelo matemático para relacionar a entrada e a saída de um processo. Eles “aprendem” essa relação a partir de um processo de “treinamento”.

4.2. Modelagem e Simulação do Processo

4.2.1. Metodologia

A etapa de modelagem através de redes neurais consiste, em geral, na determinação de pesos associados a cada entrada de dados. Estes pesos são valores que minimizam o erro quadrático entre o valor calculado pela rede e o valor experimental fornecido como saída. Sendo, então, importante ressaltar os procedimentos realizados para ajustes do modelo via redes neurais:

- a) Análise do Problema a ser modelado via redes neurais**
- b) Coleta de Dados**
- c) Análise Estatística e Reconciliação dos Dados**
- d) Verificação do domínio das variáveis da rede neural**
- e) Normalização dos dados**
- f) Treinamento da Rede Neural**
- g) Validação da rede neural treinada**

Aspectos importantes sobre estes procedimentos podem ser encontrados, de modo generalizado, no item 3 referente a Redes Neurais e serão descritos em detalhes, analisando o caso em estudo, neste capítulo.

4.2.2. Modelagem e Ajuste do Sistema

Com os dados de entrada e saída previamente selecionados, pode-se modelar as redes neurais sob diversas óticas, dentre elas: redes que modelam o processo como um todo; redes que modelam apenas partes do processo e posterior conexão entre elas a fim de modelar o sistema completo; redes que modelam de forma híbrida propriedades e partes dos processos. No caso em estudo, optou-se por dividir o processo em subsistemas. Cada subsistema representa uma grande

operação unitária (reator, coluna de destilação ou extração). Esta metodologia visa:

- Simplificar a modelagem por diminuição da ordem das redes neurais;
- Conseguir uma melhor consistência dos dados da planta dentro de cada subsistema. Erros de instrumentação em uma variável de um subsistema não prejudica, na maioria das vezes, as informações para outro subsistema;
- Uma importante parcela das correntes de reciclo do processo ficam explícitas, facilitando a análise técnico-econômica do processo.

Assim, para cada subsistema, são analisadas as variáveis medidas de entrada e saída do processo com respeito à consistência da informação e sua importância.

Entre os ganhos de informação que a etapa de modelagem das redes traz, está a detecção de erros de medição, indicando qual provável instrumento apresentou problemas em dado período.

A escolha das variáveis que representam o processo ocorreu em conjunto com o tratamento dos dados históricos utilizados no treinamento das redes. Também foi realizado o procedimento de Análise de Componentes Principais (PCA) (Khattree & Naik, 2000).

Após a modelagem das redes neurais individuais que representam os subsistemas, o foco principal é a interligação seqüencial dessas redes com a finalidade de gerar dados para construir a superfície de resposta das variáveis envolvidas, assim como analisar o comportamento das principais variáveis do processo.

Isto pode ser feito construindo uma “caixa preta”, como mostrado na Fig. 4.1 abaixo, contendo as redes interligadas pelos reciclos e correntes subseqüentes. Desta maneira, pode-se reduzir a ordem das redes neurais empregadas.

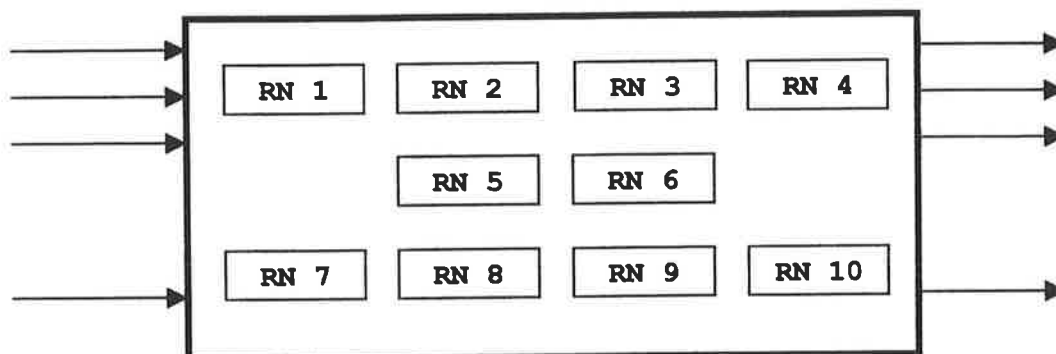


Fig. 4.1. Esquema do Modelo da Planta Industrial – “Caixa Preta”

O procedimento de convergência do sistema integrado é realizado interligando as correntes comuns aos equipamentos, reproduzindo a unidade industrial com todas as correntes modeladas pelas redes neurais. Algoritmos com métodos conhecidos foram utilizados para auxiliar a convergência matemática.

No caso estudado, inicialmente trabalhou-se com dados históricos da planta, formando um banco de dados com valores de médias diárias das variáveis de processo coletadas durante três anos de produção. Este procedimento trouxe alguns problemas no ajuste das redes neurais, por motivos de confiabilidade dos dados. Numa etapa posterior, foram utilizados dados obtidos com frequência de 15 minutos durante 01 ano de operação, o que facilitou o estabelecimento de períodos de estado estacionário e possibilitou avaliar problemas nos instrumentos de medição.

4.3. Otimização de Processos

Métodos de otimização de um dado processo envolvem minimização ou maximização de uma função objetivo que pode ser estabelecida pelo ponto de vista técnico e/ou econômico. Em geral, as variáveis de decisão estão sujeitas tanto a restrições tais como limites máximos e mínimos quanto a restrições relacionadas a considerações de segurança e àquelas que surgem das equações da

modelagem do processo. As restrições de qualidade geralmente representam especificações, enquanto as restrições de desigualdade em geral representam limites inferiores e superiores das variáveis operacionais.

Candidatos para a medida de uma função objetivo são as medidas de lucratividade, começando com medidas aproximadas, tais como taxa de retorno ou investimento, lucro do empreendimento, período de pagamento, e os custos anuais. Para uma análise completa, medidas rigorosas que envolvem o valor temporal do dinheiro e fluxos de caixa são empregadas, incluindo o valor líquido presente e a taxa de retorno do investimento. O vetor de variáveis de projeto que representam uma função objetivo é composto por variáveis que podem ser ajustadas pelo projetista e para as quais significantes mudanças são antecipadas na função objetivo. É comum começar com uma medida aproximada e ir substituindo por medidas mais rigorosas à medida que o modelo do fluxograma é refinado.

A maioria das restrições de igualdade é formada pelas equações de modelagem associadas com as unidades de processo.

A maior vantagem em executar processo de otimização é a possibilidade de incluir, explicitamente, as restrições de desigualdade, que são usados como limites da provável região de operação. No caso de não execução de otimização do processo, deve-se evitar posicionar as variáveis de projeto de modo que elas violem as restrições de desigualdade. Em alguns casos, as violações podem ser detectadas quando se examinam os resultados da simulação. Em outros casos, as sub-rotinas são incapazes de resolver as equações.

Algumas restrições de desigualdade simplesmente estabelecem limites inferiores e superiores nas variáveis. Outras permitem a especificação de um limite inferior ou um limite superior.

4.3.1. Algoritmo de Otimização

Normalmente, em problemas de engenharia química, a função objetivo e as restrições são não-lineares. Métodos computacionais de programação não linear com restrições usualmente têm que lidar com problemas tais como avaliações numéricas de derivadas. Por outro lado, métodos de derivadas livres, também chamados de métodos de busca direta, são normalmente menos eficientes e demandam maior tempo computacional, pois eles requerem um grande número de iterações. A baixa eficiência dos métodos de busca direta resulta da necessidade de resolução de modelos de equações não lineares em cada iteração. A simulação destas equações é normalmente demorada. Outra importante dificuldade em problemas de otimização é a presença de múltiplos mínimos locais, que surgem a partir de várias corridas do programa de otimização com diferentes pontos iniciais das variáveis de decisão.

O caminho mais fácil e direto para melhorar a função objetivo é por repetição da simulação. Neste procedimento, o projetista seleciona valores das variáveis de projeto e completa a simulação. Então, em geral, usando uma estratégia sistemática, ajusta-se as variáveis de projeto e repete-se a simulação usando, por exemplo, análise de sensibilidade nos simuladores de processos, nos quais os resultados da simulação são recalculados à medida que as variáveis de projeto são ajustadas usando incrementos uniformes entre limites especificados pelo usuário. Entretanto, análise de sensibilidade pode ser computacionalmente lenta e gerar extensos arquivos de informação, muitos dos quais estão associados a processos sub ótimos. Alternativamente, o projetista pode selecionar um algoritmo de otimização construído no simulador para ajustar as variáveis de projeto, uma estratégia que, em geral, é eficiente. Entretanto, quando muitos ciclos estão presentes, os cálculos de simulação podem consumir muito tempo, com cerca de 20 ou mais iterações requeridas para atingir um ponto ótimo e, para cada uma destas iterações, são necessárias 20 ou mais iterações para convergir cada um dos ciclos e circuitos de controle (envolvendo especificações de projeto). Para superar estas ineficiências, especialmente para fluxos integrados, a

última estratégia ajusta as variáveis de projeto e as variáveis de corte para convergir todos os ciclos simultaneamente (Seider et al., 1999).

Na abordagem de simulações repetidas, partindo de uma estimativa inicial para as variáveis, a simulação é completada quando os ciclos convergem; isto é, as equações de corte são satisfeitas. Então, a função objetiva é ajustada, até certo grau, arbitrariamente, e a simulação é repetida usando a solução prévia como estimativa inicial. Esta estratégia é repetida até atingir a convergência para um mínimo.

Este procedimento é lento e demorado, o que torna atrativo o uso de Redes Neurais como ferramenta de otimização, uma vez que lidam facilmente com restrições. Usar Redes Neurais, para resolver problemas de otimização, implica no levantamento total da superfície de resposta do processo com restrições. A partir dessa superfície pode-se chegar aos mínimos globais das variáveis ou até mesmo determinar a região ótima de operação. Alternativas recentes para procedimentos de otimização baseada no uso de redes neurais torna possível a busca detalhada do “grid search”. Esta nova aproximação evita vários problemas numéricos típicos das ferramentas convencionais de otimização, além de apresentar maior velocidade de processamento. Entretanto, o sucesso na obtenção de uma rede neural robusta e apropriada depende fortemente da escolha das variáveis de processo envolvidas, bem como da qualidade dos dados disponíveis e do domínio usado para fins de treinamento. Se qualquer comportamento importante do processo não for coberto por observações, o modelo de redes neurais ou qualquer outro modelo empírico terá uma representação mais pobre e a otimização provavelmente falhará.

4.3.2. Abordagem para Otimização com Base em Redes Neurais

A idéia principal é substituir as equações de modelagem por uma rede neural equivalente, empregando dados históricos da planta industrial e utilizar esta rede

para executar um “grid” de busca detalhado da região de interesse, através um mapeamento completo da função objetivo no espaço das variáveis de decisão. Usar redes neurais em substituição ao modelo fenomenológico apresenta a vantagem de resposta mais rápida da simulação via redes neurais. Desta forma, mesmo que o “grid” detalhado possa ser atingido em tempo razoável, desde que não existam muitas variáveis sendo otimizadas (neste caso, surgem problemas de dimensionalidade). Esta abordagem é mais confiável, lida com restrições, evita vários problemas numéricos típicos das ferramentas convencionais de otimização e apresenta maior velocidade computacional. Como benefício adicional, o mapeamento completo da função objetivo permite identificar facilmente múltiplos pontos ótimos, uma importante característica não apresentada pelos métodos convencionais de otimização. Além disso, as restrições são facilmente tratadas desde que pontos com restrições violadas possam ser reconhecidos e classificados de acordo com a importância da restrição.

Uma vez que o mapa é obtido, é fácil escolher o ponto ótimo, identificar se múltiplos ótimos estão presentes, checar se restrições foram violadas, entre outros. Esta abordagem definitivamente fornece mais informações para análise do engenheiro que os procedimentos convencionais de programação não linear. Este algoritmo pode ser estendido diretamente para tratar problemas de otimização multi-objetiva.

4.3.3. Metodologia

A otimização pode promover melhorias econômicas e/ou técnicas operacionais. Problemas envolvendo otimização apresentam etapas típicas que devem ser resolvidas, sendo a mais importante, a etapa de formulação do problema. Então, torna-se necessário identificar os elementos essenciais que descrevem o sistema a ser estudado, organizando-os matematicamente da seguinte forma:

- Estabelecer o modelo do processo (restrições e comportamento do sistema): O modelo e as restrições descrevem a influência das variáveis no sistema e como podem afetar os objetivos desejados.

- Estabelecer a função objetivo (econômica): A função objetivo representa custo, lucro, energia, rendimentos, condições operacionais, entre outros, em termos das variáveis do processo em estudo.

O desenvolvimento de um modelo confiável que represente bem o processo é de grande importância para que os resultados da otimização tenham coerência e sejam reais. Os resultados da modelagem da Unidade de Produção de Isopreno via redes neurais mostrou-se bom para esta aplicação já que os resultados gerados pelo modelo das principais variáveis representaram bem o comportamento dos dados da planta industrial.

Em certos casos, o uso de métodos matemáticos de otimização não traz benefícios, por exemplo, quando uma solução razoável pode ser obtida baseada em experiências passadas, quando existe uma prática padrão em uso, quando o tempo necessário para avaliar o problema não é compatível com o contexto do projeto global e quando as informações necessárias podem ser obtidas com grande custo.

No caso em estudo, o modelo foi desenvolvido via redes neurais devido, principalmente, à grande dificuldade na modelagem fenomenológica da Seção de Extração. Assim, a otimização do comportamento das variáveis deste processo é possível utilizando experiências passadas e informações adicionais como restrições, sem a necessidade do conhecimento físico-químico do processo.

A metodologia empregada na otimização do processo é a apresentada por Nascimento & Giudici (1998). O emprego da técnica de redes neurais para este tipo de problema apresenta a vantagem de se poder varrer toda a solução do

processo mudando-se as variáveis do processo dentro da precisão da instrumentação. Como o número de variáveis é, em geral, grande, o número possível de combinações chega a casa de milhões. O fato do modelo de redes neurais ser computacionalmente muito simples faz com que milhões de casos possam ser calculados em minutos (o que ainda hoje não seria viável, em termos de tempo, empregando modelos fenomenológicos nesta busca exaustiva da solução). Com a descrição de toda a superfície de resposta o problema de mínimos locais deixa de existir (problema esse dos métodos convencionais de otimização) e as soluções, que na maioria das vezes não são muitas, são encontradas.

De modo geral, uma seqüência de etapas para a resolução de problemas de otimização é apresentada a seguir:

a) Análise do processo:

Analisar o processo de forma a definir as variáveis e características específicas de interesse, bem como seus limites comportamentais.

b) Desenvolvimento do modelo do processo:

Desenvolver através de expressões matemáticas ou relações empíricas um modelo representativo do processo, identificando as variáveis dependentes e independentes, além das restrições de igualdade e desigualdade. As restrições de qualidade do produto, de segurança e econômicas impostas são relações de desigualdade.

c) Determinação da função objetivo específica e o critério de otimização

O procedimento empregado na otimização não requer necessariamente uma função objetiva formal, podendo-se obter tanto otimizações qualitativas (a partir de conjunto de restrições de processo e de análise de ganhos econômicos) como quantitativas (com função objetiva econômica).

Deve-se ter em mente, nesta etapa, qual o principal objetivo a ser atendido, por exemplo, maior produção, dada especificação de produto. Assim, a função objetivo escolhida pode atender apenas aspectos operacionais, como por exemplo, especificar quais modificações podem ser efetuadas em variáveis como vazões de reciclo, vapor para refulvedor, temperaturas, para determinados valores de vazão e composição de carga,.

d) Definição matemática do processo

Esta etapa envolve a escolha de uma técnica de otimização adequada para definir matematicamente o processo em estudo. Como o problema trata da otimização não linear com restrições e o interesse é demonstrar a capacidade do modelo em fornecer algumas condições operacionais ótimas dentro de alguns limites factíveis, a opção é pelo método de busca direta.

4.4. Aplicação à Planta de ISOPRENO

A planta industrial estudada é a unidade de produção de isopreno da BRASKEM.

4.4.1. Descrição do Processo

A unidade de extração de isopreno tem por objetivo produzir isopreno de alta pureza para obtenção de borracha sintética e borracha termoplástica, a partir de um corrente C5 de gasolina de pirólise.

A separação do isopreno é realizada a partir do processo de destilação extrativa usando como solvente uma mistura de acetonitrila (ACN), água (H₂O), álcool alílico (AAL) e acetona em proporções especificadas. O processo pode ser dividido, basicamente, em 5 sistemas principais: preparação de carga ou sistema de concentração de Isopreno, extração, recuperação de solvente, remoção de enxofre e fracionamento.

A Fig. 4.2. mostra, esquematicamente, a Unidade de Produção de Isopreno e cada seção é descrita a seguir. As diferentes cores das correntes de processo têm apenas o objetivo de facilitar a visualização do fluxo de processo: As linhas em vermelho indicam o fluxo seguido pelo componente principal, o isopreno; as linhas em rosa representam os subprodutos e o fluxo do refinado; a cor azul indica o circuito do solvente e a cor verde, o circuito de água.

4.4.1.1. Sistema de Concentração de Isopreno

O Sistema de Concentração de Isopreno tem por finalidade preparar a carga para as demais seções, isto é, separar da fração C5 de gasolina de pirólise, proveniente de outras unidades da BRASKEM, uma corrente rica em isopreno (IP), com um teor máximo de 4,0% de ciclopentadieno (CPD).

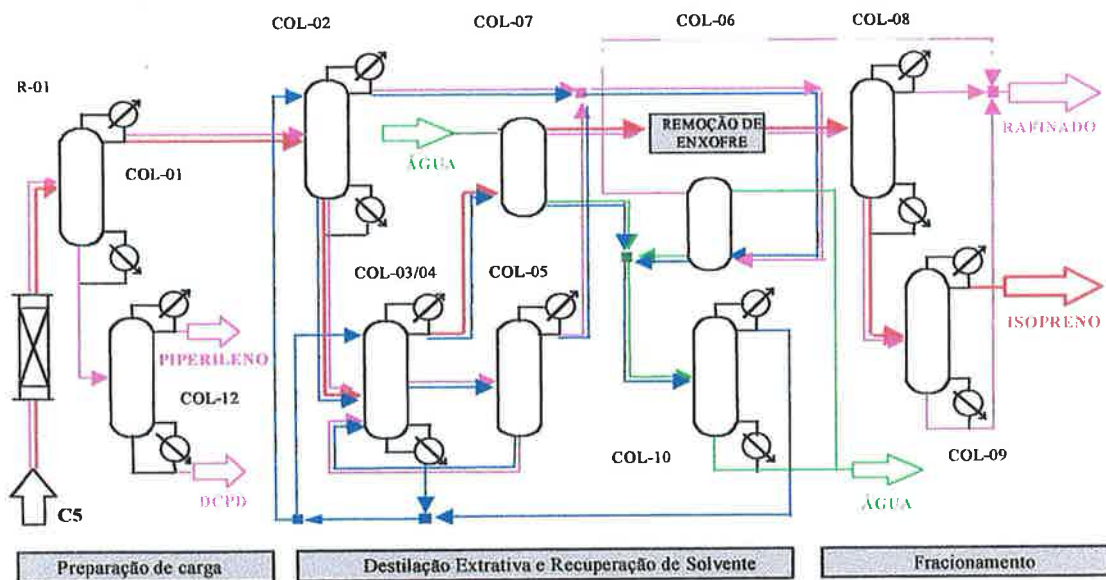


Fig. 4.2. Fluxograma de Processo da Unidade de Isopreno

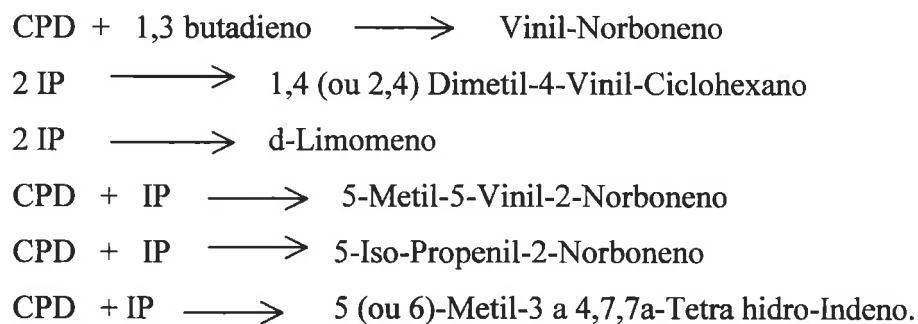
O sistema consta de um reator (R-01) onde o CPD é dimerizado a dicitopentadieno, DCPD, uma coluna superfracionadora (COL-01A/B) tendo como produto de topo a carga para a extração e produto de fundo os compostos C10 formados no reator e parte do CPD não dimerizado e uma coluna depentanizadora (COL-12) que separa a corrente de fundo da COL-01B em uma corrente C5 rica em piperilenos e uma corrente C10 rica em DCPD.

No Reator de Dimerização (R-01), a carga é submetida a um tratamento térmico, no qual ocorrem reações de dimerização que consomem CPD e isopreno. O sistema reativo é muito complexo e, uma simplificação das reações que ocorrem são apresentadas a seguir:

Reação principal - Dimerização do CPD a DCPD:



Reações paralelas:



A corrente rica em piperilenos é comercializada com um teor mínimo de n-pentano de 1%. Para tanto a COL-01 A/B tem a função de separar o n-pentano dos piperilenos. Para manter a especificação do CPD na carga da Extração foi aumentado o tempo de estocagem da carga para possibilitar uma maior dimerização de CPD, tornando possível o ajuste final no reator.

Pelas equações de balanço de massa e energia, o perfil de temperatura do reator apresenta uma relação definida com a composição inicial da alimentação e conversão no reator. Isto leva a um controle de temperatura da carga do reator.

A superfracionadora, devido ao grande número de pratos necessários à separação, é dividida em duas seções. A corrente efluente do reator passa através da válvula de controle de pressão do reator, onde é parcialmente vaporizada.

Parte do vapor de topo é retirado como produto, sendo enviado para o Sistema de Extração de Isopreno. O restante passa através de um trocador, onde é totalmente condensado com água de refrigeração, sendo recolhido no vaso de refluxo, o qual possui uma bota para retirada de água. O efluente deste é, então, enviado, como refluxo, para coluna COL-01 A. O controlador de vazão do produto de topo é ajustado de forma a especificar o n-pentano no topo da COL-12 quando da produção de Piperilenos.

A retirada de produto líquido de fundo da COL-01B é enviada, por diferença de pressão, para a coluna COL-12. O vapor de topo da COL-12 é totalmente

condensado e direcionado ao vaso de acúmulo. Parte do condensado retorna à coluna como refluxo e parte é retirado como produto de topo, uma corrente rica em piperilenos, a qual é enviada para o vaso de refinado ou para estocagem na esfera de Piperilenos.

O produto de fundo da COL-12, uma corrente rica em DCPD, é transferida para estocagem em tanque (como produto final ou para queima), sendo antes resfriado com água de resfriamento.

4.4.1.2. Sistema de Extração e Lavagem

A fração de C5 proveniente da seção de dimerização é alimentada na fase vapor ao topo da COL-02 B (seção de fundo da Primeira Coluna de Destilação Extrativa).

A destilação extrativa usando acetonitrila como solvente ocorre na COL-02 A/B e o refinado contendo pentano, pentenos e pentadieno-1,4 é descarregado pelo topo da coluna (COL-02 A).

O solvente rico contendo isopreno, componentes de volatilidade menor que o isopreno, como o ciclopentadieno, piperilenos e acetilenos e uma pequena quantidade de pentenos, é enviada do fundo da coluna COL-02 B à Coluna Stripper de Solvente (COL-04).

A corrente rica em solvente, contendo ciclopentadieno, piperilenos, acetilenos e isopreno, é enviada do fundo da Segunda Coluna de Destilação Extrativa (COL-03) ao topo da COL-04.

O solvente é separado dos hidrocarbonetos na Coluna Stripper de Solvente (COL-04), e o solvente pobre, completamente livre de hidrocarbonetos, é descarregado pelo fundo da coluna.

Na seção abaixo da bandeja de alimentação da COL-04, isopreno é recuperado do solvente juntamente com parte do ciclopentadieno, piperilenos e acetilenos alimentados. Estes três últimos, como apresentam uma volatilidade relativa menor que o isopreno, são em parte concentrados num estágio intermediário e retirados como uma corrente lateral para a COL-05.

O vapor, contendo isopreno, ciclopentadieno, piperilenos e acetilenos, é enviado do topo da COL-04 para o fundo da Segunda Coluna de Destilação Extrativa (COL-03), sem ser condensado.

A destilação extrativa, usando acetonitrila como solvente, ocorre na COL-03, e o isopreno é, então, parcialmente separado dos componentes que apresentam menores volatilidades relativas, tais como ciclopentadieno, piperilenos e acetilenos. Esta separação é conseguida operando-se a COL-03 com uma relação solvente/hidrocarbonetos menor que a usada na COL-02 A/B, de modo a concentrar o vapor produto de topo no componente de maior volatilidade relativa, que é o isopreno.

Isopreno cru é descarregado do topo da COL-03 e enviado à coluna lavadora de produtos (COL-07), onde o solvente, que acompanha o isopreno cru (por azeotropia), é removido e enviado ao sistema de remoção de enxofre.

O solvente rico, contendo ciclopentadieno e isopreno, é enviado do fundo da COL-03 de volta a COL-04.

Visto que a corrente lateral descarregada da COL-04 é acompanhada de hidrocarbonetos (principalmente pentadienos e acetilenos) e também vapor de solvente, ela é enviada à Coluna de Remoção de CPD (COL-05) e separada em solvente e hidrocarbonetos. Os hidrocarbonetos retirados pelo topo são condensados e enviados à coluna lavadora de C5 (COL-06). O líquido no fundo da coluna de remoção de CPD (COL-05) retorna à COL-04.

4.4.1.3. Circulação de Solvente e Sistema de Regeneração

O solvente pobre do fundo da Coluna Stripper de Solvente (COL-04) circula para as Colunas de Destilação Extrativa (COL-02A e COL-03), após seu calor sensível ter sido recuperado e utilizado. Visto que dímeros podem ser formados e acumulados nesta circulação do solvente durante a operação, uma parte do solvente é descarregada continuamente como purga e o dímero separado para regenerar o solvente.

O refinado separado do topo da Primeira Coluna de Destilação Extrativa (COL-02A) é dividido em duas correntes. A maior parte da corrente é enviada à Coluna Lavadora de C5 (COL-06), juntamente com o destilado do topo da Coluna de Remoção de CPD (COL-05). A água, separada da bota do vaso de refluxo da Primeira Coluna de Destilação Extrativa, e a corrente de purga de solvente são enviadas ao Vaso Separador de Solvente (VS-07) como descrito a seguir:

O solvente que acompanha o refinado separado do topo da Primeira Coluna de Destilação Extrativa (COL-02A), e uma pequena parte de solvente, ainda contido no destilado de topo da coluna de remoção de CPD (COL-05), são separados por lavagem com água na coluna lavadora de C5 (COL-06). Os hidrocarbonetos são enviados do topo da COL-06 para o vaso de refinado C5 (VR-14), e então enviados para o limite de bateria.

Uma parte da corrente de purga de solvente é misturada a uma corrente de refinado do topo da Primeira Coluna de Destilação Extrativa (COL-02A), como mencionado, e enviada à Coluna Lavadora de C5 (COL-06). O restante da corrente de purga de solvente é misturado a uma corrente de refinado desviada do topo da Primeira Coluna de Destilação Extrativa (COL-02A) e água da bota do vaso de refluxo da COL-02A, através do misturador estático e enviado ao Vaso Separador de Solvente (VS-07). Os hidrocarbonetos em fase líquida são separados da água, que contém solvente, no VS-07.

O dímero na corrente de purga de solvente é transferido para o líquido leve, e este enviado à Coluna Lavadora de C5 (COL-06). Água mais solvente são retornados ao sistema de circulação de solvente.

Água contendo solvente é descarregada da Coluna Lavadora de Produto (COL-07) e Coluna Lavadora de C5 (COL-06). Estas correntes são misturadas e enviadas à Primeira Coluna de Recuperação de Solvente (COL-10).

O solvente é recuperado pelo topo da COL-10 por destilação e retornado como solvente circulante. Uma parte da água do fundo da coluna é enviada à Coluna Lavadora de C5 (COL-06) e o restante descarregado do sistema.

O calor latente do vapor de topo COL-10 pode ser usado como fonte de calor para o refulvedor da Primeira Coluna de Destilação Extrativa (COL-02A/B).

4.4.1.4. Sistema de Remoção de Enxofre

Isopreno cru (topo da COL-07) contendo enxofre, principalmente na forma de CS₂, oriundo da carga, é alimentado a um sistema de filtragem para remoção de pequenas partículas sólidas e em seguida a um coalescedor para remoção de água livre. Posteriormente esta corrente é enviada a um vaso contendo 2 leitos de peneira molecular tipo 3A, para remoção de água solúvel. A corrente resultante, com menor teor de água, é alimentada a outro vaso contendo 2 leitos de alumina específica para remoção, por adsorção, de enxofre. Isopreno cru, com o enxofre especificado, é enviado para o sistema de fracionamento, na coluna COL-08.

A redução do teor de água é necessária pois a alumina também adsorve água concorrendo com o enxofre e diminuindo o tempo de campanha. Os leitos de peneira e alumina se saturam, passando a operar o outro leito previamente reativado. Este sistema não é considerado neste estudo.

4.4.1.5. Sistema de Fracionamento

A purificação do Isopreno é realizada em duas colunas de fracionamento: uma removedora de leves (COL-08), e uma segunda removedora de pesados (COL-09), obtendo-se no topo desta última o isopreno produto com pureza acima de 99,9%.

Isopreno cru, proveniente do sistema de recuperação de enxofre, é alimentado à Coluna de Remoção de Leves (COL-08), onde butino-2 e pentadieno-1,4 são separados por destilação e descarregados pelo topo, sendo, então, enviados ao vaso de refinado C5 (VR-14).

O fundo da COL-08 é enviado à Segunda Coluna de Remoção de Pesados (COL-09), onde ciclopentadieno (CPD) e outros componentes da fração pesada são separados e descarregados pelo fundo. O produto, isopreno, é obtido pelo topo da torre e bombeado para o resfriador de isopreno, sendo então enviado para o vaso de isopreno produto (shift tank). A partir deste ponto o isopreno é enviado para o limite de bateria.

4.4.1.6. Separação de Álcool e Acetona

A fração C5 contém compostos com oxigênio (acetona, álcoois, etc), e alguns deles (álcool alílico e acetona, por exemplo), podem se acumular gradativamente no solvente circulante durante a operação. Um acúmulo excessivo destes compostos pode afetar a seletividade do solvente e por conseguinte a destilação extrativa, sendo necessário removê-los na Segunda Coluna de Recuperação de Solvente (COL-11).

Na operação de remoção do álcool alílico no solvente circulante, o líquido do fundo da Primeira Coluna de Recuperação de Solvente (COL-10) é enviado a

COL-11. O álcool alílico é separado pelo topo da coluna e enviado ao vaso de refinado C5 (VR-14).

Uma parte da água de fundo da coluna é enviada à Coluna Lavadora de C5 (COL-06) e o restante descarregado do sistema. A separação de álcool alílico na COL-11 deve ser feita periódica e intermitentemente.

A COL-11 também pode ser usada na remoção de acetona do solvente circulante, quando não estiver na operação de álcool alílico. O vapor do solvente descarregado do mais baixo estágio da Coluna Stripper de Solvente (COL-04) é enviado a COL-11 para destilação. A acetona é separada pelo topo e enviada ao vaso de refinado C5 (VR-14), o solvente é enviado do fundo da coluna para o fundo da COL-04, para recuperação. A COL-11 não é considerada neste estudo

Como dito previamente, para efeito de facilitar a modelagem usando redes neurais o processo foi dividido em 10 subsistemas, cada um dos quais será apresentado e discutido detalhadamente nas seções seguintes. Os 10 subsistemas são:

- **Subsistema1 (RN1)** – Reator (R-01)
- **Subsistema2 (RN2)** – Primeira Coluna de Remoção de Pesados (COL-01A/B)
- **Subsistema 3 (RN3)** – Coluna Depentanizadora (COL-12)
- **Subsistema 4 (RN4)** – Primeira Coluna de Destilação Extrativa (COL-02A/B)
- **Subsistema 5 (RN5)** – Segunda Coluna de Destilação Extrativa (COL-03) e Coluna Stripper de Solvente (COL-04)

- **Subsistema 6 (RN6)** – Coluna de Remoção de CPD (COL-05)
- **Subsistema 7 (RN7)** – Coluna Lavadora de C5 (COL-06)
e Coluna Lavadora de Produto (COL-07)
- **Subsistema 8 (RN8)** – Primeira Coluna de Recuperação de Solvente (COL-10)
- **Subsistema 9 (RN9)** – Coluna de Remoção de Leves (COL-08)
- **Subsistema 10 (RN10)** – Segunda Coluna de Remoção de Pesados (COL-09)

4.4.2. Metodologia

A estrutura global do modelo de otimização, seguindo as etapas previamente descritas, é mostrada na Fig. 4.3.

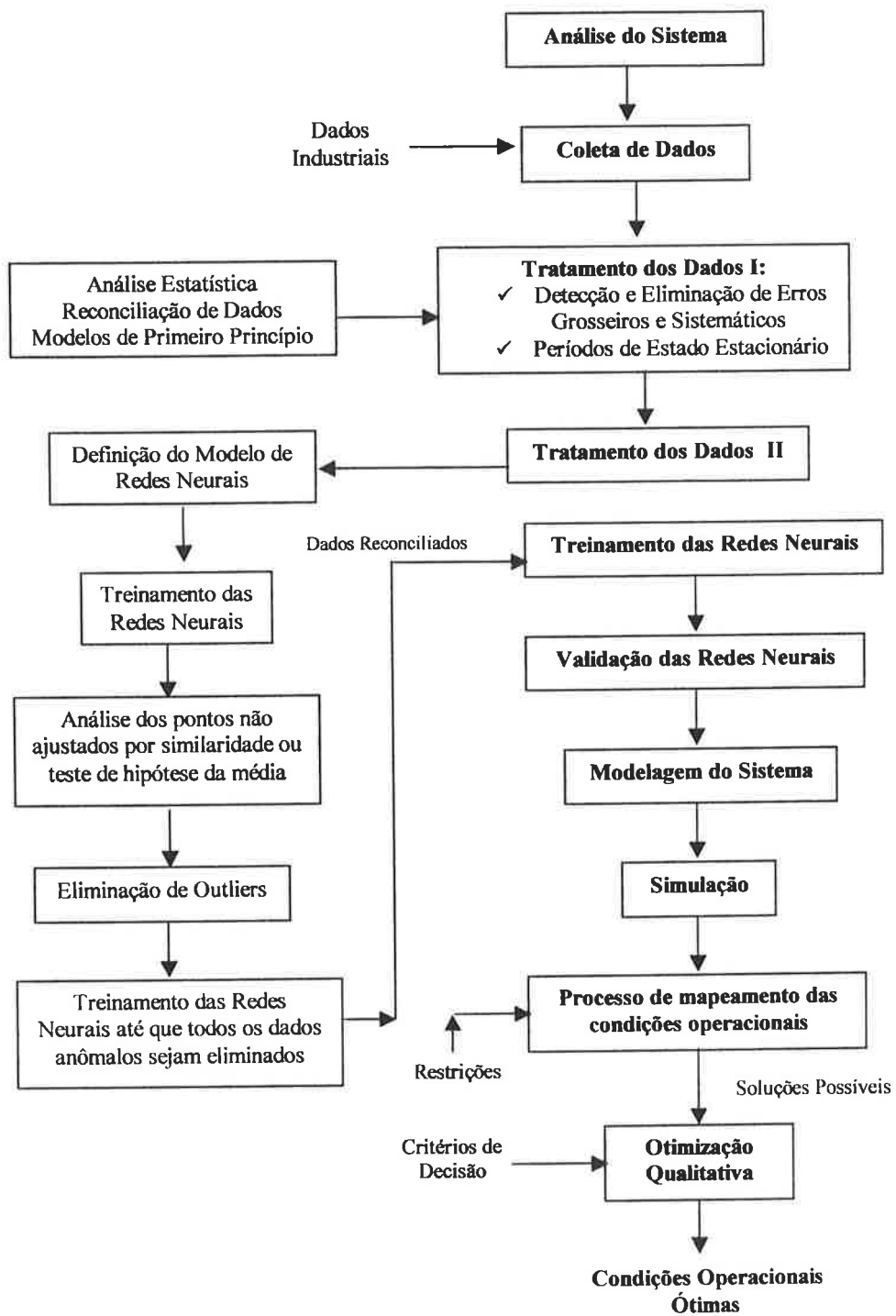


Fig. 4.3. Esquema do Procedimento de Otimização

4.4.2.1. Dados Industriais

Foram utilizados dois conjuntos distintos de dados históricos da unidade de Isopreno. O primeiro é constituído por médias diárias das principais variáveis do processo obtidas durante os anos de 1997, 1998 e 1999. A análise de reconciliação de dados foi executada para cada ano individualmente. O segundo é um banco dados das principais variáveis de processo obtidas durante um ano com frequência de 15 minutos. Este novo conjunto é mais representativo do processo, pois problemas de amostragem em períodos curtos do dia podem, evidentemente, ser incorporados na média diária, não sendo possível detectá-los. Nos casos de erros de instrumentação em períodos mais longos, a média reflete esse erro. Desta forma, é possível identificar períodos de operação em estado estacionário, além de detectar problemas nos instrumentos de medida.

O banco de dados usado na modelagem da planta industrial é, portanto, apresentado contendo dados das principais variáveis do processo durante um ano de operação, correspondendo ao período de setembro de 2000 a agosto de 2001. O banco de dados básico consiste de cerca de 34500 conjuntos de observações de 244 variáveis.

4.4.2.2. Tratamento dos Dados

A modelagem usando redes neurais foi feita diretamente dos dados industriais. A presença de ruídos e erros grosseiros e sistemáticos, comumente encontrados em processos industriais, pode causar distúrbios na modelagem e conseqüente perda de qualidade do ajuste. Portanto, para este trabalho, a primeira etapa foi preparar os dados industriais para o ajuste do modelo através da detecção e eliminação dos dados contendo erros grosseiros e sistemáticos.

Esta fase do trabalho é de extrema importância já que a qualidade dos dados influi diretamente na qualidade do ajuste realizado pelas redes neurais. São

utilizados, nesta fase, além do conhecimento do processo, procedimentos estatísticos e outros baseados nas equações do primeiro princípio.

Este tratamento foi dividido em partes, a saber:

1. Seleção das Variáveis de Interesse
2. Identificação de Erros Grosseiros
3. Identificação de Erros Sistemáticos
4. Estabelecimento de Períodos de Operação em Estado Estacionário
5. Redução do número de conjunto de dados através cálculo da média para os períodos de estado estacionário
6. Tratamento dos dados reduzidos utilizando redes neurais

De acordo com o esquema da Fig. 4.3, esta divisão foi considerada da seguinte forma:

Parte I:

- **Seleção das variáveis de interesse**

As variáveis de interesse foram definidas considerando os dados de processo disponíveis e sua importância para o processo e operação da planta. Então, valores mínimos, máximos e médios foram identificados, bem como a variância para cada variável. Variáveis cuja faixa operacional encontrava-se muito próximo dos limites dos instrumentos de medição (fundo de escala) não foram incluídas na análise, bem como aquelas para as quais a variância era muito pequena (valores praticamente constantes).

Nesta etapa também se fez uso da metodologia de Análise de Componentes Principais (PCA) de modo a selecionar o melhor conjunto de variáveis de entrada e saída para cada subsistema a ser modelado via redes neurais. Uma vez mais o conhecimento do processo foi de grande importância, já que em alguns casos os

resultados obtidos através desta análise não foram propriamente os esperados de acordo com a importância de certas variáveis sabidamente relevantes para o processo e/ou operação da planta. Neste caso, optou-se em utilizar as variáveis que atendiam o pré-requisito exigido pelo processo. Explicações para algumas destas diferenças já têm sido encontradas, enquanto para outras, ainda fazem-se necessários alguns esclarecimentos sobre o processo e operação da planta. Para este trabalho, utilizou-se o programa SAS conforme descrito em Khattree & Naik, 2000.

- **Identificação de erros grosseiros**

Erros grosseiros ocorrem, possivelmente, devido a alterações nas condições de operação previstas ou não, erros na operação de instrumentos de medidas, erros no registro da informação, o que pode contaminar dados válidos. “Outliers” também podem ser simplesmente um dos valores extremos em uma distribuição de probabilidade que ocorre naturalmente, mas não freqüentemente e não deve ou não pode ser rejeitado. Nesta etapa, também foram avaliados e eliminados os seguintes dados: valores nulos, negativos, valores que apresentam diferentes ordens de grandeza e descontinuidade brusca no comportamento das variáveis em função do tempo. Esta análise foi feita através de gráficos das variáveis de interesse em função do tempo e verificação de seus valores mínimos, máximos e médios, observando também a variância dos dados. A experiência do processo permitiu que alguns pontos fossem eliminados por estarem fora de uma faixa aceitável ou possível de operação.

- **Identificação de erros grosseiros e/ou sistemáticos**

Quando se conhece a causa do valor irregular, pode-se sem hesitação descartar tal observação (vide item anterior). No entanto, quando não se tem certeza do erro, ou não se tem prática suficiente em aceitar ou rejeitar uma

observação extrema, deve-se basear o julgamento em algum tipo de análise estatística e/ou procedimentos que utilizam equações do primeiro princípio. Experiência tanto no aspecto estatístico como do processo produtivo também é importante para avaliação de tais pontos.

Uma vez efetuados os balanços materiais globais e por componentes, foi possível identificar algumas distorções nos resultados finais. Neste ponto foram muito significantes as informações obtidas através da análise da representação gráfica de uma distribuição de frequência. Verificando se as distribuições apresentavam forma normal, pôde-se avaliar o número de pontos em termos percentuais do número total de observações que se encontravam abaixo ou acima de um determinado valor. Embora distribuições de dados reais, devido ao tamanho finito do conjunto, nunca possam ser perfeitamente normais na forma, a aproximação é geralmente próxima o suficiente para permitir usar a distribuição normal teórica como um modelo para interpretar populações empíricas de dados. Para identificar a localização relativa de um valor observado em uma dada distribuição é necessário conhecer a média aritmética do conjunto de dados, seu desvio da média, e como estes desvios podem ser expressos como desvios padrão. Uma simples inspeção dos resultados finais pode ser útil na detecção, eliminação ou correção de valores estranhos ao conjunto de dados. Para tanto, tomou-se por base o fato de ser sabido que a adição ou subtração de um valor constante de um conjunto de observações afeta apenas a média destes valores, mas não a variação dos dados, enquanto a multiplicação ou divisão por uma constante afeta tanto a média quanto à variação da distribuição original (Kachigan, 1991).

Nesta etapa também foi feita a comparação com dados de análise de laboratório, mais confiáveis. Estes procedimentos permitiram identificar possíveis erros de instrumentos de medida em determinados períodos de operação e conseqüente correção dos valores provavelmente incorretos, usando a metodologia apresentada no parágrafo anterior.

- **Identificação de períodos de operação em estado estacionário**

A maior frequência na obtenção dos dados permitiu a identificação de períodos de operação em regime estacionário. O principal critério adotado foi vazão de alimentação permanecer constante por um período de dois a três dias. Uma flutuação nos dados de 0,2 a 0,3 t/h foi aceitável.

Posteriormente, também foi verificada a constância ou variação nas demais variáveis envolvidas, adotando-se como critério faixas de valores dentro de um intervalo de confiança estabelecido.

Parte II:

- **Redução do conjunto de dados**

O conjunto de dados resultante da eliminação de erros grosseiros e sistemáticos foi reduzido a um novo conjunto constituído pelas médias dos períodos identificados com base nos itens anteriores.

- **Tratamento dos dados utilizando redes neurais**

A partir dos dados provenientes da etapa anterior, foi feito o treinamento das redes neurais representativos de cada subsistema, utilizando o conjunto completo de dados, e em seguida foram verificados o erro residual do ajuste e o gráfico do valor experimental da variável em função do seu valor calculado pela rede. Dessa forma pode-se identificar pontos discrepantes em relação à tendência da maioria dos dados. Em seguida, esse ponto foi analisado em relação ao significado no processo e mediante técnicas estatísticas pôde-se decidir pela rejeição ou não do mesmo. Foi empregada a metodologia de Análise de “Clusters” em conjunto com o princípio da similaridade ou teste de hipótese de médias. O procedimento foi

então repetido até que não houvesse mais pontos anômalos e a precisão desejada fosse atingida.

Esta metodologia é lenta e trabalhosa, pois cada ponto considerado irregular requer análise individual cada vez que a rede é treinada. Espera-se que com a experiência adquirida neste estudo, possa-se estabelecer um procedimento sistemático para executar tal análise em outras situações. Outra alternativa seria utilizar a técnica de Análise de Componentes Principais para a detecção de “outliers”, conforme descrito em Khattree & Naik (2000).

4.4.2.3. Modelagem

Depois da análise dos dados da planta, o processo de modelagem foi executado. Esta etapa consistiu em estabelecer uma estrutura de rede neural ótima para representar cada unidade de processo individualmente, isto é, determinar o número de neurônios na camada oculta, o número de iterações para o processo de aprendizagem, o valor residual e selecionar os conjuntos de treino e teste.

Inicialmente, os dados industriais foram divididos em dois grupos: um conjunto de treinamento e um conjunto de teste. O conjunto de teste continha dados dentro da faixa de domínio dos dados de treinamento. Assim, não ocorrem extrapolações no modelo. Para a divisão dos conjuntos foi utilizado o programa PREP2.FOR desenvolvido por Guardani (1997)

Então, as diversas redes neurais foram executadas, individualmente, várias vezes, com diferentes especificações, isto é, variando o número de neurônios na camada oculta (NH) e o número de apresentações (iterações), de modo a selecionar a rede mais adequada para representar o sistema. Deste modo várias redes com diferentes erros foram gerados para cada subsistema. Para evitar sobre-ajuste do modelo, escolheu-se o menor número de neurônios na camada

oculta, bem como um apropriado número de ciclos de treinamento (iterações) que corresponde ao menor erro no conjunto de teste. A fig. 4.4. mostra um exemplo deste procedimento, isto é, o mínimo erro para o conjunto de teste como função do número de neurônios e apresentações. O valor mínimo para o erro é atingido para NH igual a 9 e cerca de 30000 apresentações podem ser observadas. O erro é representado pelo quadrado da diferença entre os valores experimentais e os calculados pela RN, conforme a eq. (3.1):

$$\min(E) = \min \left[\sum_{m=1}^r \sum_{k=1}^p (y_k^{(m)} - O_k^{(m)})^2 \right] \quad (3.1)$$

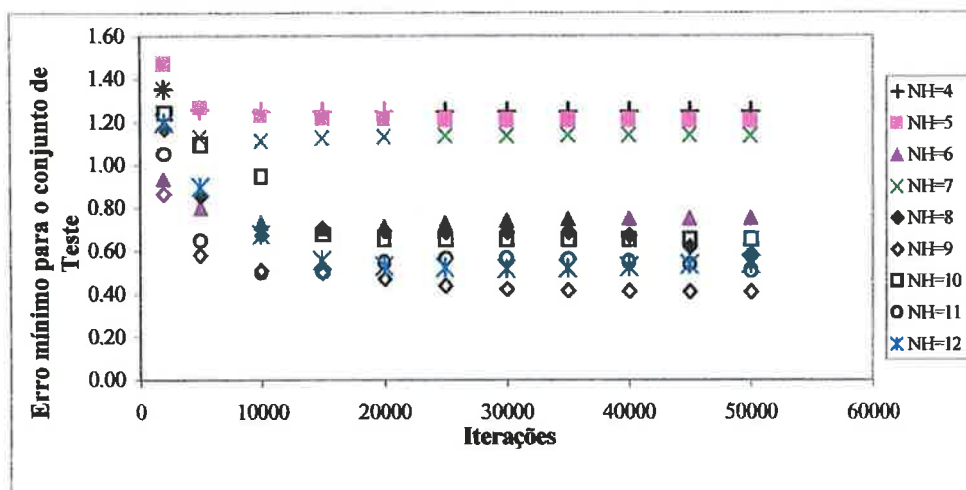


Fig. 4.4. Erro Mínimo para o Conjunto de Teste - Exemplo.

Após a modelagem das redes neurais individuais, as quais representam cada subsistema, o principal foco foi interconectar seqüencialmente estas redes de modo a gerar dados com o objetivo de construir uma superfície de resposta para todas as variáveis envolvidas, bem como analisar o comportamento das variáveis de processos mais importantes. Assim uma “caixa preta” foi construída, como mostrado no exemplo da fig. 4.1, contendo todas as redes interconectadas por correntes seqüenciais e recíclós.

O procedimento de convergência do sistema integrado foi executado interconectando todas as correntes seqüencialmente às correspondentes operações unitárias ou subsistemas, de modo a reproduzir a planta industrial. O algoritmo de substituição direta com fator de amortecimento foi usado para convergência matemática.

O modelo global é complexo devido ao grande número de variáveis e ciclos envolvidos, no entanto tanto o modelo gerado e a estrutura de convergência empregada mostraram boa eficiência na representação do sistema.

4.4.2.4. Otimização

A primeira etapa foi fornecer as especificações das correntes de interesse, por exemplo, vazão total e composição. Então, o mapeamento da solução foi iniciado mudando-se os valores das variáveis de processo e, nesta etapa, apenas as soluções que satisfizeram determinadas condições operacionais foram aceitas. Estas condições podem ser composição de componentes importantes para o processo, perda, temperaturas, etc., as quais também devem estar em uma dada faixa.

A etapa seguinte consistiu da escolha qualitativa das condições operacionais que poderiam minimizar os custos de energia; estas condições podem ser a menor vazão de refluxo, a menor vazão de vapor para o refeedor e a menor temperatura de carga. Claro que satisfazer todas estas condições muitas vezes não é possível, e é, então, necessário fazer uso de algumas regras hierárquicas.

4.5. Resultados e Discussões

4.5.1. Preparação dos Dados

A abordagem do problema de análise de dados irregulares depende do objetivo em questão: deseja-se apenas determinar se uma observação é um “outlier” de modo a investigar as condições que podem ter levado a esta observação extrema, ou se o interesse é apenas eliminar tais pontos de modo a obter um conjunto mais uniforme. Neste estudo, deseja-se inicialmente detectar e eliminar os dados anômalos e, em uma segunda etapa, tentar explicar as possíveis causas de erro.

Uma primeira análise dos dados foi realizada de acordo com seguintes critérios, conforme definidos no item 4.4.2.2.- Parte I, para eliminação de valores irregulares: valores completamente fora de uma faixa aceitável para cada variável; observação de gráficos das variáveis analisadas em função do tempo, considerando os possíveis patamares surgidos e descontinuidade dos pontos; experiência tanto no aspecto estatístico como do processo produtivo; balanço material e de energia dos equipamentos. Os resultados para os dados obtidos a cada 15 min podem ser vistos nas Figs. 4.5. a 4.7 para o subsistema 2 (COL-01), como exemplo da metodologia empregada. Não serão apresentados os resultados de todo o procedimento para cada subsistema, uma vez que poderia tornar-se cansativa a visão de uma série de gráficos similares para a maioria das 244 variáveis envolvidas no sistema. Casos especiais e relevantes serão tratados quando da modelagem de cada subsistema, quando, após a etapa de tratamento, foram selecionadas as variáveis de maior importância para o processo e operação da unidade.

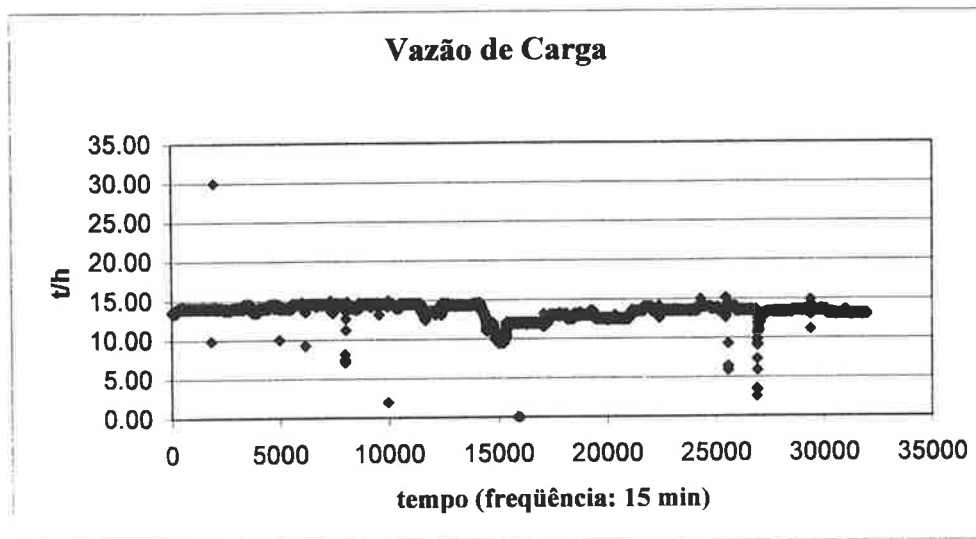


Fig. 4.5. Conjunto de Dados Experimentais Originais

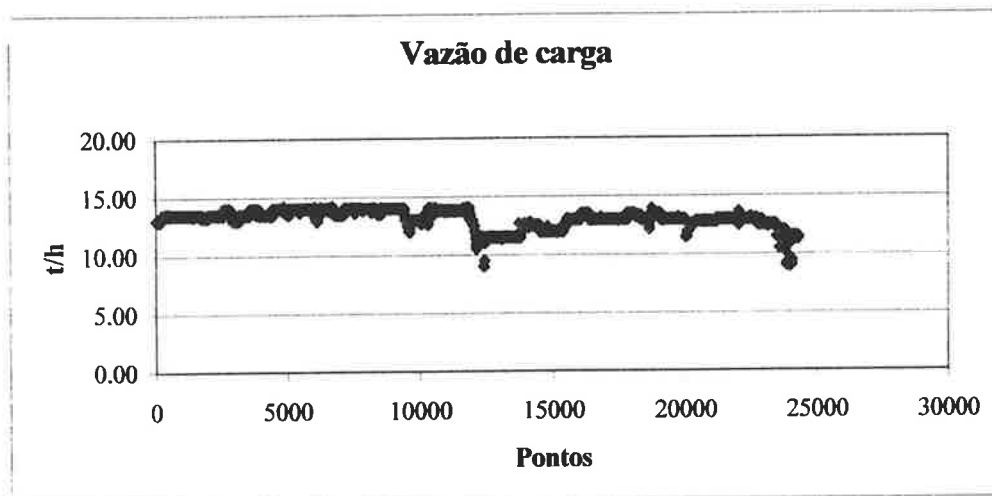


Fig. 4.6. Conjunto de Dados Experimentais Após Eliminação de "Outliers"

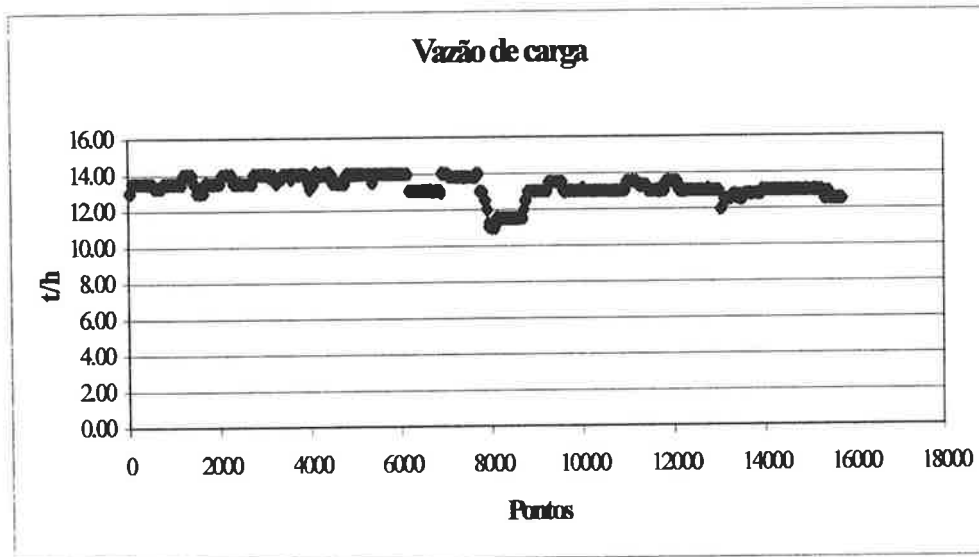


Fig. 4.7. Conjunto de Dados Experimentais Após Eliminação de “Outliers” e Erros Sistemáticos

Detecção de “outliers” ou erros grosseiros não se constituiu uma tarefa difícil de ser realizada devido à frequência de coleta dos dados. No caso de erros sistemáticos, no entanto, a tarefa foi um pouco mais complexa e demorada, primeiro, devido à necessidade de determinar todos os balanços globais e por componentes para cada subsistema individualmente, estabelecer períodos de regime em estado estacionário, calcular medias e desvios padrão e em seguida construir histogramas para cada um dos balanços em cada período estacionário. A análise destes diagramas e resultados permitiu detectar erros sistemáticos e eliminá-los ou corrigi-los de acordo com a disponibilidade de dados de laboratório para comparação, empregando a metodologia apresentada no item 4.4.2.2.- Parte I. Os resultados podem ser vistos nas Fig. 4.8 a 4.11.

A Fig. 4.8 mostra o histograma antes da análise para o balanço material global do subsistema 2 (COL-01).

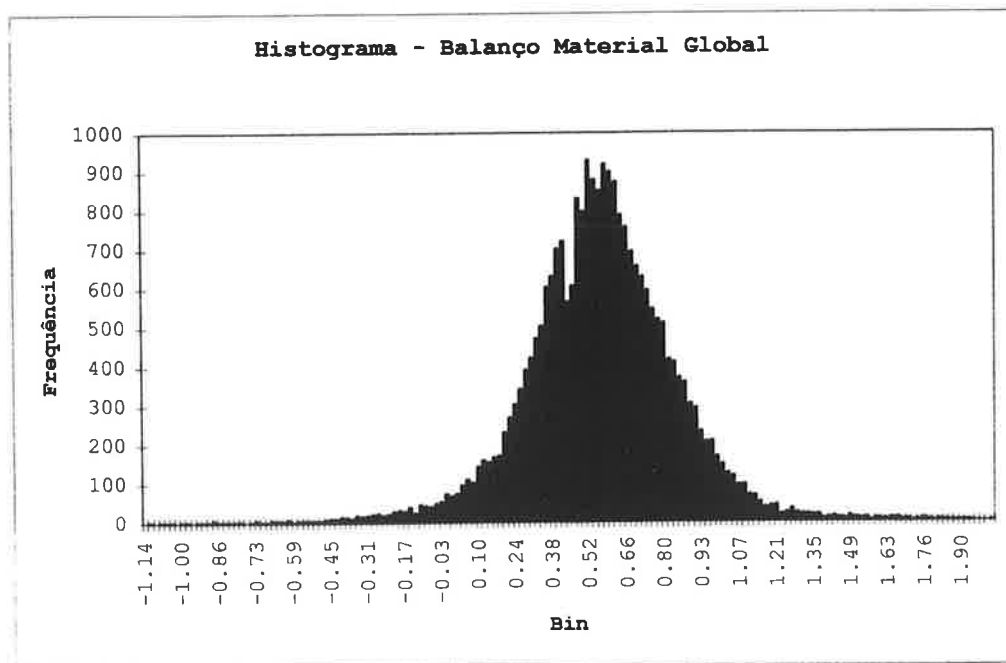


Fig. 4.8. - Balanço Material Global Antes do Ajuste e Eliminação de Erros Sistemáticos

Nesta figura, pode ser visto que a média é igual a 0,55, o que significa que uma vez que a distribuição é normal, o único problema é o deslocamento da média e baseado no fato de que o balanço material global deve ser igual a zero, foi possível corrigir tal desvio. Então analisando os balanços de massa global e por componentes tinham-se duas opções: diminuir a vazão de alimentação ou aumentar as vazões de topo e/ou fundo. Uma vez que as vazões de alimentação apresentam valores mais altos, a decisão foi pela primeira opção, isto é, diminuir a vazão de alimentação através de subtração do valor da média. O histograma final tem a mesma forma do original, mas com média igual a zero, como pode ser visto na Fig. 4.9.

As Figs. 4.10 e 4.11 mostram os resultados antes e depois da análise do balanço material de Isopreno ao redor do subsistema 2 (COL-01). Neste caso a análise foi realizada comparando os dados experimentais da planta com os dados obtidos através de ensaios de laboratório. Para a maioria dos pontos, foi possível corrigir os valores multiplicando ou dividindo o valor da planta por um dado fator obtido pelo balanço teórico do componente, considerando as vazões experimentais corretas.

Neste caso, verifica-se que o histograma final apresenta uma forma mais próxima da distribuição normal, embora a média seja apenas aproximadamente igual a zero. Isto pode ser explicado pelo fato que se conhecem, apenas, as medições experimentais das análises em linha de Isopreno para as correntes de alimentação e topo e, sabe-se, pelas volatilidades relativas dos componentes na mistura, que praticamente todo Isopreno é obtido no topo da coluna.

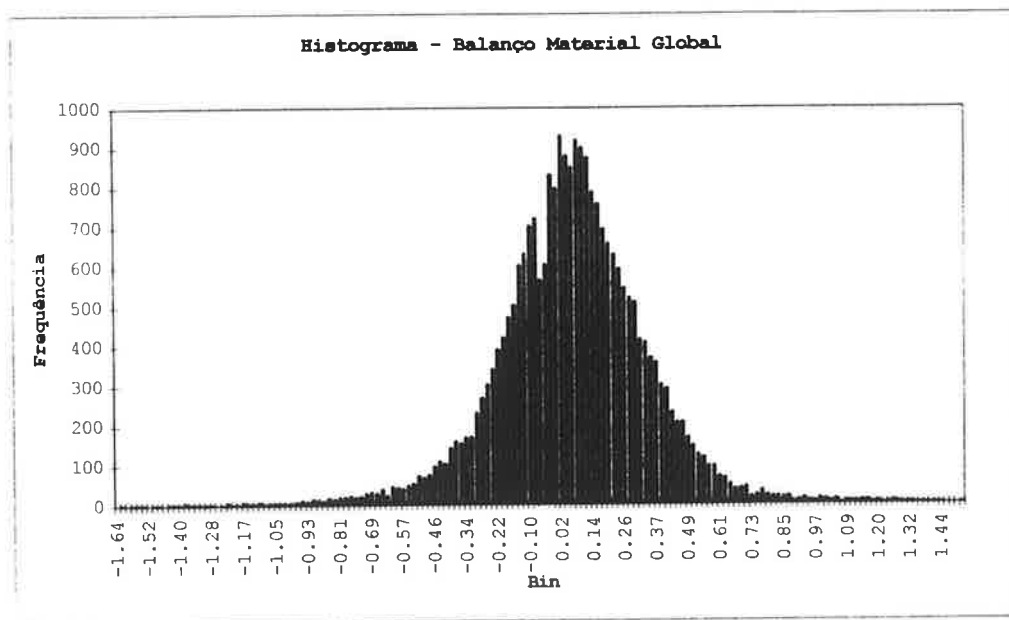


Fig. 4.9. - Balanço Material Global Após Ajuste e Eliminação de Erros Sistemáticos

Após esta análise inicial, o resultante conjunto de dados, após redução do número de conjuntos de variáveis, foi utilizado em uma primeira fase do treinamento da rede neural de modo a proceder a segunda etapa de reconciliação de dados, conforme descrito na Parte 2 do item 4.4.2.2. Para tal, todos os dados foram incluídos no conjunto de aprendizagem e a partir do resultado da rede treinada, pôde-se perceber que alguns pontos não foram ajustados adequadamente, o que é uma indicação de possíveis problemas de consistência ou erros não detectados na primeira etapa do estudo.

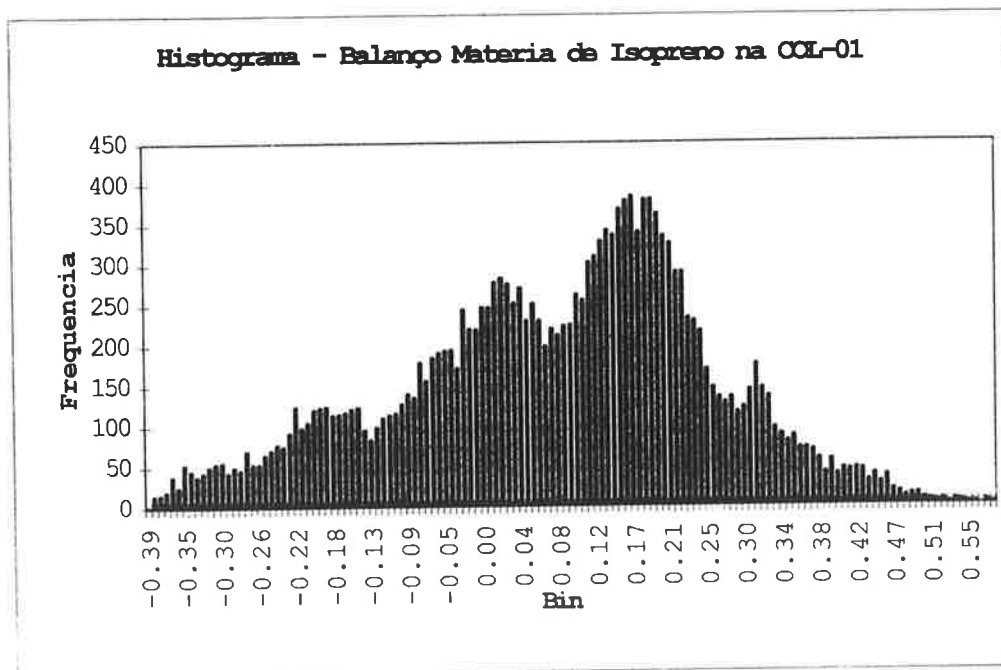


Fig. 4.10. Histograma do Balanço de IP antes da correção e eliminação de erros sistemáticos

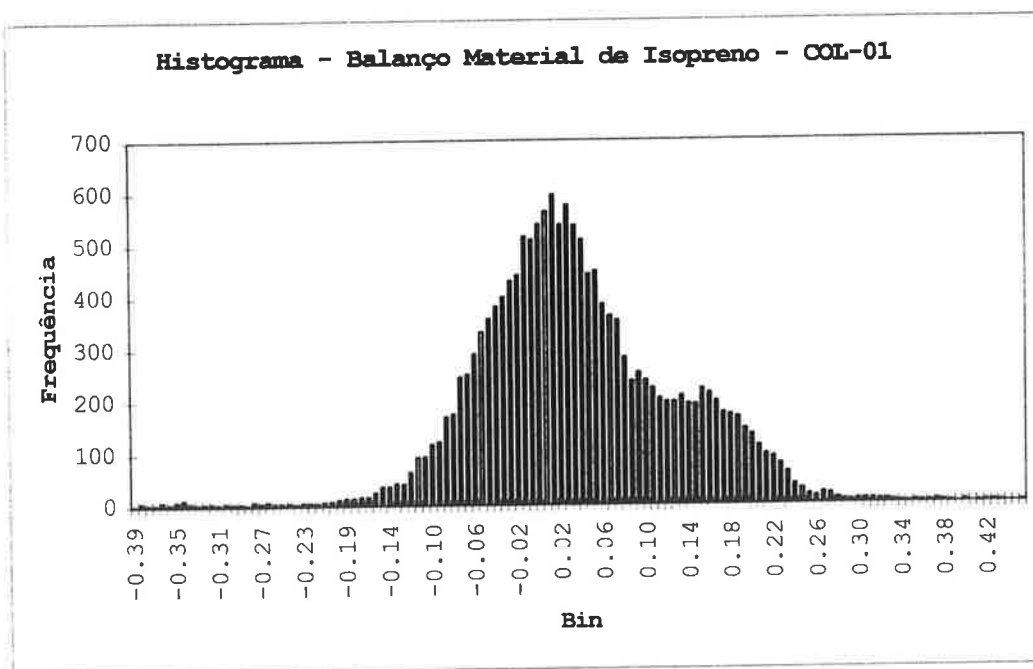


Figura 4.11. Histograma do Balanço de IP após correção e eliminação de erros sistemáticos

De modo a verificar a consistência destes dados e sua conseqüente possível eliminação, os mesmos foram analisados com base no princípio da similaridade entre vários conjuntos de dados (análise de grupos). Considerando que um determinado conjunto de dados é um vetor constituído pelas variáveis selecionadas de entrada (ENTRADA) e de saída (SAÍDA) para cada subsistema ou operação unitária, de modo que cada conjunto de dados corresponde à informação de um dia de operação, no caso de trabalhar-se com dados de média diária, ou valor médio da variável obtido após redução do número de conjunto de dados, no caso de trabalhar com dados obtidos com freqüência de 15 min. É de se esperar que para uma série de ENTRADAS semelhantes, as variáveis de saída do processo (as variáveis dependentes) apresentem valores semelhantes. Quando se observa uma dada SAÍDA ou uma dada ENTRADA diferente entre uma série de semelhantes, o conjunto de dados correspondente pode ser eliminado. Exemplos podem ser vistos na Tabela 4.1.

Tabela 4.1. Análise de Grupos - Exemplos

Data	Entradas						Saídas						
	Ent1	Ent2	Ent3	Ent4	Ent5	Ent6	Ent7	Ent8	Sai1	Sai2	Sai3	Sai4	
Ago/97	22	12,79	15,34	1,86	63,80	39,30	59,90	8,89	4,37	8,35	25,24	2,50	97,4
Ago/97	23	12,80	15,40	1,88	63,80	39,20	59,90	8,87	4,35	8,35	27,21	2,63	97,3
Ago/97	24	12,80	15,25	1,79	63,80	39,20	59,90	8,85	4,38	8,35	26,88	2,53	97,4
Ago/97	25	12,75	15,07	1,71	63,80	39,10	59,80	8,79	4,36	8,35	27,21	2,49	97,3
Out/97	04	13,99	16,03	2,40	64,50	55,10	59,80	11,19	4,83	9,08	20,85	2,68	102,6
Dez/97	17	13,82	16,15	2,93	64,30	55,50	59,30	10,79	5,06	9,16	28,84	2,89	97,2
Dez/97	18	13,80	16,09	2,91	64,30	4,36	59,40	10,81	4,97	9,23	28,57	2,85	97,8
Dez/97	29	13,70	16,02	2,74	64,20	55,60	58,60	11,13	4,91	9,07	30,50	2,66	95,0
Dez/97	30	13,70	15,86	2,54	64,20	55,80	58,60	11,15	5,07	8,92	30,92	2,54	94,6

Em alguns casos, não é possível uma análise de forma direta, por exemplo, quando um determinado conjunto de dados é único, não sendo, portanto, possível comparação ou quando existem apenas dois conjuntos para comparação, não sendo possível determinar qual das informações pode estar incorreta. Para estes casos, foi utilizado, o teste de hipótese das médias. O teste das médias compreende a definição do intervalo de confiança e teste de hipótese. Uma vez estabelecido o intervalo de confiança para uma determinada amostra do conjunto

total de dados, verifica-se se o conjunto de dados problemático satisfaz a hipótese nula, isto é se o valor em questão está incluído no intervalo de confiança, então o dado não pode ser eliminado; caso contrário, se o dado estiver fora dos limites do intervalo de confiança, este é rejeitado. Para este estudo, foi admitido um coeficiente de confiança de 95%. Exemplos podem ser vistos na Tabela 4.2. a seguir.

Tabela 4.2. Análise de Teste de Hipótese das Médias – Exemplos

Data		Entradas						Saídas					
		Ent1	Ent2	Ent3	Ent4	Ent5	Ent6	Ent7	Ent8	Sai1	Sai2	Sai3	Sai4
Jul/97	14	13,00	14,50	1,95	66,00	51,10	59,90	10,51	5,55	7,62	26,73	1,62	93.5
Jul/97	20	13,00	14,57	2,77	65,90	50,60	60,00	10,25	5,37	7,72	26,15	2,68	91.6
Set/97	10	13,20	16,20	2,43	64,30	50,40	59,70	9,94	4,41	8,68	25,81	3,17	100.1
Jun/97	29	12,00	13,58	2,38	66,00	50,20	61,20	10,40	3,77	8,78	19,91	2,77	97.2
Jun/97	30	12,00	14,74	2,14	65,90	48,40	61,30	10,08	3,77	8,89	20,77	2,70	98.5
Jun/97	01	12,09	14,13	1,75	66,50	48,60	61,00	10,23	4,04	8,58	23,51	2,07	97.5
Jul/97	04	13,30	14,99	2,46	66,80	48,80	60,90	10,32	5,15	8,73	24,35	2,70	94.8
Jul/97	05	13,48	15,74	2,37	66,80	49,60	60,70	10,47	5,35	8,76	24,60	2,47	93.8
Jul/97	06	13,50	15,64	2,23	66,80	50,30	60,80	10,57	5,48	8,72	25,22	2,30	93.8
Jul/97	07	13,51	15,02	2,01	66,70	50,50	60,80	10,56	5,51	8,76	25,23	2,26	94.0
Jul/97	08	13,50	15,87	1,96	66,40	49,10	60,60	10,34	5,49	8,79	25,91	2,17	94.5
Jul/97	14	13,00	14,50	1,95	66,00	51,10	59,90	10,51	5,55	7,62	26,73	1,62	93.5
Jul/97	18	12,98	13,52	2,22	65,70	51,60	60,00	10,65	5,38	7,74	26,09	2,57	93.1
Jul/97	20	13,00	14,57	2,77	65,90	50,60	60,00	10,25	5,37	7,72	26,15	2,68	91.6
Set/97	10	13,20	16,20	2,43	64,30	50,40	59,70	9,94	4,41	8,68	25,81	3,17	100.1
Set/97	12	13,00	14,06	2,05	64,30	51,20	59,70	10,27	4,28	8,65	20,93	2,66	100.5
Set/97	13	13,38	14,13	1,98	64,30	50,10	59,40	10,38	4,41	8,84	20,98	2,52	100.8
Set/97	16	14,01	13,99	2,60	64,50	52,30	59,60	10,97	4,40	9,45	17,98	2,94	100.9
Set/97	17	14,00	14,24	2,64	64,40	52,30	59,50	11,00	4,61	9,30	18,49	2,89	99.6
Set/97	18	14,00	14,11	2,57	64,40	52,30	59,50	11,00	4,62	9,30	19,09	2,89	99.9
Set/97	19	14,01	15,67	2,59	64,30	50,70	59,70	10,81	4,56	9,41	19,65	2,82	99.8
Mínimo		12,00	13,52	1,75	64,30	48,40	59,40	9,94	3,77	7,62	17,98	1,62	91.60
Máximo		14,01	16,20	2,77	66,80	52,30	61,30	11,00	5,55	9,45	26,73	3,17	100.9
Média		13,22	14,71	2,28	65,56	50,45	60,24	10,49	4,79	8,71	22,86	2,57	96.88
Desvpad		0,65	0,82	0,29	1,04	1,23	0,66	0,31	0,63	0,54	3,06	0,37	3.20
t(n-1)=18	2,11												
méd+desvpad*t		14.60	16,44	2,90	67,75	53,05	61,63	11,13	6,10	9,84	29,31	3,35	103,6
méd-desvpad*t		11.84	12,97	1,67	63,37	47,85	58,85	9,84	3,47	7,57	16,41	1,78	90,13

Uma simples medida para avaliar a qualidade do ajuste da rede neural escolhida para os dados experimentais é geralmente a comparação entre os valores calculados pela rede neural com o dados experimentais. A dispersão dos pontos de dados ao redor da linha ideal de 45° pode ser usada para julgar o ajuste da rede neural aos dados experimentais. A idéia para usar redes neurais com o objetivo de detecção de outliers é baseada neste tipo de diagrama (Büllau et al., 1999). Por consequência, pode-se mostrar que os “outliers” dos dados experimentais correspondem aos “outliers” desta curva. Seguindo esta metodologia, a rede neural foi primeiramente treinada para o conjunto completo de dados e depois para cada conjunto filtrado. Os “outliers” detectados após a primeira corrida de treinamento foram analisados através dos métodos estatísticos descritos previamente. Este procedimento foi então repetido diversas vezes até que os dados não mais apresentassem pontos anômalos. Desde que o treinamento de cada rede após os dados filtrados leva a diferentes resultados, a base de dados de entrada varia devido ao procedimento de filtração ou eliminação de “outliers”. A Tabela 4.3. mostra um exemplo do número de pontos eliminados em cada corrida e as Figs. 4.12.a-c. apresentam alguns exemplos dos resultados deste método, para a primeira, segunda e última corridas, respectivamente.

Tabela 4.3. Pontos eliminados em cada corrida
Exemplo

Corrida	Número de pontos eliminados
1	49
2	28
3	19
4	7
5	12
6	-

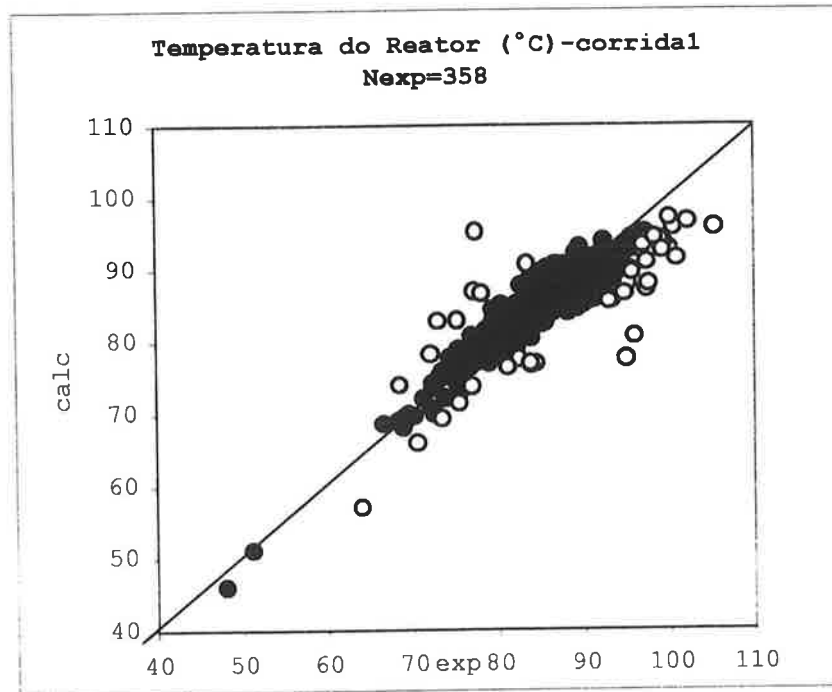


Fig. 4.12.a. Comparação entre valores calculados e dados medidos: antes da análise

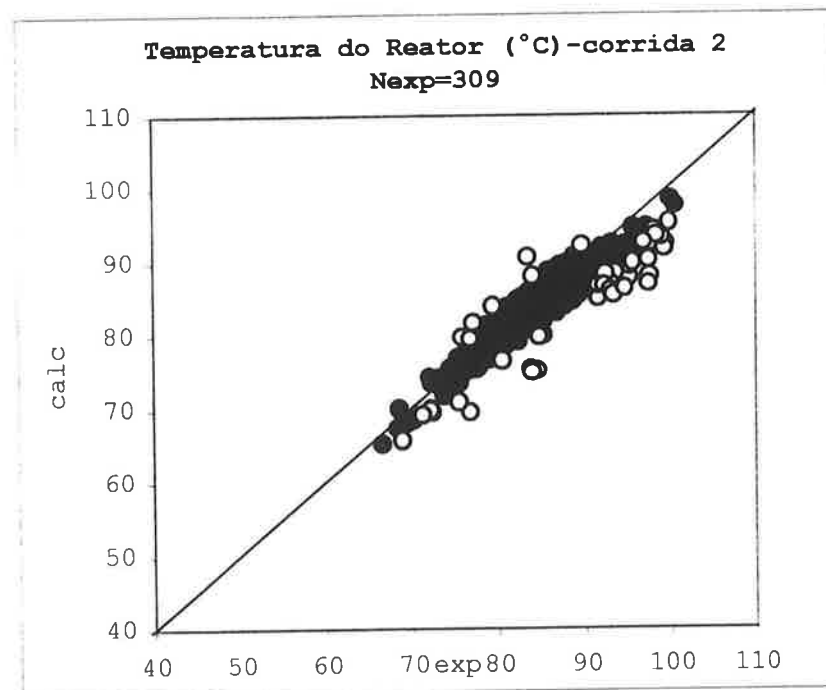


Fig. 4.12.b. Comparação entre valores calculados e dados medidos: resultados intermediários;

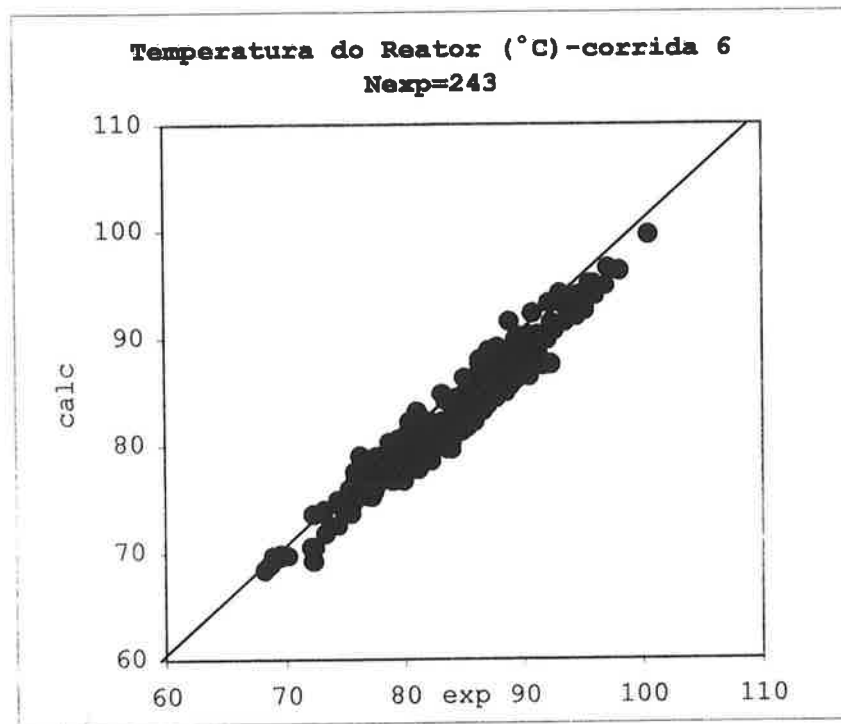


Fig. 4.12.c. Comparação entre valores calculados e dados medidos:
resultado final

4.5.2. Modelagem

Conforme descrita anteriormente no item 4.2.2, a metodologia adotada para a modelagem do processo foi baseada na divisão do processo em subsistemas e posterior conexão entre eles. A partir deste ponto, serão apresentados os resultados detalhados obtidos no estudo individual de cada subsistema, incluindo tratamento dos dados, observações, comentários, treinamento, ajustes e seleção da configuração da rede neural mais adequada a cada caso. Em todos os casos, inicialmente foram selecionadas as variáveis de processo mais importante entre as disponíveis, bem como a sua faixa de aplicabilidade. Foi utilizado, como ferramenta auxiliar na etapa de seleção das variáveis, o método de Análise de Componentes Principais (PCA), de modo a direcionar o julgamento. No entanto, a decisão final seguiu as diretrizes básicas obtidas junto ao corpo técnico da BRASKEM quanto às variáveis críticas para o processo.

Neste trabalho, para treinamento da rede foi utilizado o programa NEURO14, desenvolvido no Laboratório de Simulação e Controle de Processos do Departamento de Engenharia Química/Escola Politécnica da USP (Nascimento, 1992).

A rede treinada é do tipo “feedforward”, com algoritmo de retropropagação (“backpropagation”) para convergência.

4.5.2.1. Subsistema 1 – REDE 1

A Fig. 4.13 apresenta o esquema de fluxo do Reator a ser modelado pela REDE1, enquanto que a rede representativa deste subsistema é apresentada esquematicamente na Fig. 4.14, mostrando as variáveis de entrada e saída selecionadas. As faixas de aplicabilidade destas variáveis podem ser vistas na Tabela 4.4.

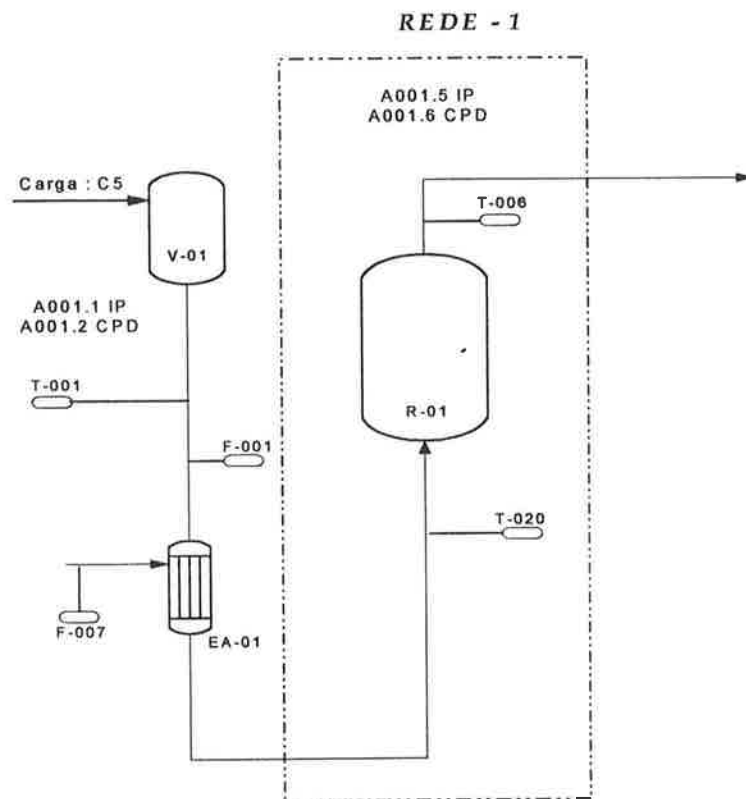


Fig. 4.13. – Fluxograma do reator modelado pela REDE 1

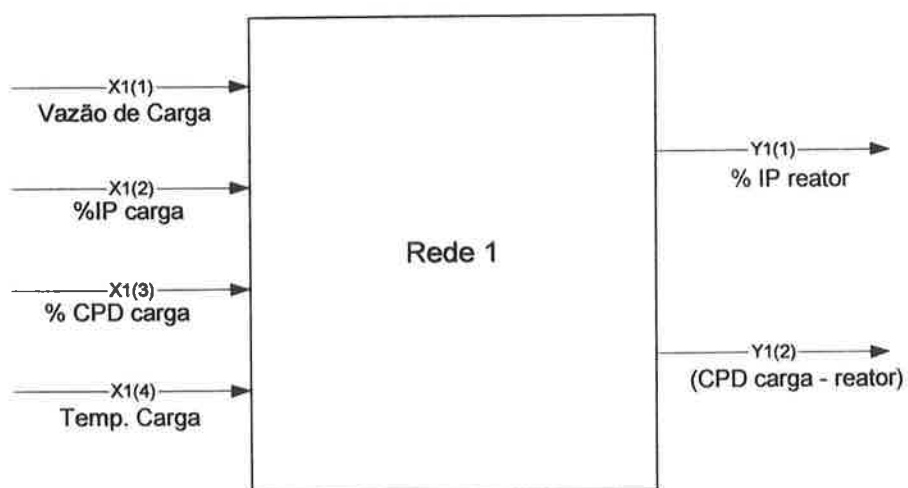


Fig. 4.14. - Esquema da REDE 1 – variáveis de entrada/saída

Tabela 4.4. - Variáveis de Processo para a REDE 1 e Suas Faixas de Aplicação

Notação	Variáveis	Unidades	Valores mínimos	Valores máximos
Entrada				
x1(1)	Vazão de carga	t/h	9,50	14,52
x1(2)	Análise de IP na carga	%massa	11,57	20,86
x1(3)	Análise de CPD na carga	%massa	2,31	19,59
x1(4)	Temperatura da Carga	(°C)	55,79	80,18
Saída				
y1(1)	Análise de IP no reator	%massa	9,80	20,24
Y1(2)	CPD carga - CPD reator	%massa	0,20	17,45

Notas: IP – Isopreno; CPD - Ciclopentadieno

De acordo com a metodologia apresentada, uma vez definidas as variáveis de entrada e saída, foi realizada uma primeira etapa de treinamento, contendo todos os pontos no conjunto de aprendizagem, de modo a verificar a consistência dos dados e eliminação de dados anômalos como descrito anteriormente na Parte II do item 4.4.2.2. Em uma segunda etapa, os dados foram divididos em dois grupos: conjunto de aprendizagem e conjunto de testes, de modo a selecionar a configuração de rede mais adequada à melhor representação do sistema. Para tanto foram feitas diversas corridas variando o número de neurônios na camada oculta e os resultados da variação do erro residual do conjunto de teste em função do número de iterações e neurônios na camada oculta é apresentado na Fig. A.1 (ANEXO A). O erro é representado pelo quadrado da diferença entre os valores experimentais e os calculados pela RN, conforme a eq. (3.1):

$$\min(E) = \min \left[\sum_{m=1}^r \sum_{k=1}^p (y_k^{(m)} - O_k^{(m)})^2 \right]$$

O número ótimo de neurônios na camada oculta (NH) corresponde ao menor erro obtido para o conjunto de testes. Assim, o número ideal de neurônio na camada oculta seria 14. Além disso, pode-se também verificar a influência do número de iterações no erro residual. Observa-se que aproximadamente acima de 40000 iterações o ganho na diminuição do erro não é significativo o suficiente para justificar a utilização de um maior número de iterações. Por outro lado, como regra heurística, é aconselhável escolher a rede neural com o menor número possível de neurônios na camada oculta e com o menor número possível de iterações que apresentem resultados satisfatórios. Isto diminui a possibilidade de ocorrência de sobreajuste, o que é um problema comum no uso de redes neurais (Pollard et al., 1992). Assim, embora o número de neurônios na camada oculta seja o parâmetro mais importante na configuração da rede, quanto maior esse valor, maior também é o número de pesos sinápticos ajustados. Diante disto, verificou-se que, neste caso, pode-se usar uma rede com 7 neurônios na camada interna, sem perda de qualidade no desempenho da representação do sistema, como será visto nos testes de validação e consistência dos resultados obtidos.

A topologia ou configurações da rede neural selecionada é apresentada na Tabela 4.5 para a REDE 1.

Tabela 4.5. - Configurações da REDE 1

REDE 1	
Número de Pontos do Conjunto de Aprendizagem	171
Número de Pontos do Conjunto de teste	79
Número de Variáveis de Entrada	4
Número de Variáveis de Saída	2
Número de Neurônios na Camada Oculta	7
Número de Neurônios na Camada de Saída	2
Número de Iterações	40000
Fator de Amortecimento	1,00
Parâmetro de Momento	1,00

A etapa seguinte foi verificar o ajuste do modelo da rede neural selecionada comparando os valores das variáveis de saída medidas e calculadas. As figuras,

mostrando os resultados obtidos no ajuste da rede, para o conjunto de dados de teste, são apresentadas no ANEXO B.

Boa Concordância entre os valores modelados e medidos podem ser observados e a distribuição uniforme indica ausência de qualquer tendência no ajuste.

O segundo teste de consistência diz respeito à distribuição da função distância, definida como a diferença entre os valores experimentais e calculados, $(y_{\text{exp}} - y_{\text{calc}})_{RN}$, ou análise da dispersão dos valores calculados das variáveis de saída em comparação aos valores medidos. Estas informações fornecem a distribuição da diferença do erro absoluto em cada ponto e avaliam a boa qualidade do ajuste. Tanto melhor o ajuste quanto mais eqüitativa a distribuição. As figuras apresentando os gráficos da Função Distância para cada variável de saída, para o conjunto de dados de teste são encontradas no ANEXO B.

Pode-se perceber, pela simetria, que o modelo apresenta um bom ajuste para ambas as variáveis de saída.

Outra forma de apresentação é a Distribuição Normal dos erros entre os valores calculados pela rede neural e os experimentais. A distribuição normal demonstra o equilíbrio de frequência dos valores encontrados em torno da média zero. Uma distribuição normal mostra a boa qualidade do ajuste realizado. Os histogramas da distribuição de erros para as variáveis de saída da REDE 1 são apresentados no ANEXO B, para o conjunto de dados de teste.

A análise dos histogramas indica que o ajuste é muito bom, já que a distribuição dos erros apresenta um comportamento normal com média aproximadamente igual a zero.

De modo geral, os erros avaliados pelas técnicas mostradas não apresentam uma variabilidade apreciável, evidenciando, portanto, um bom ajuste do modelo

da REDE 1. Este bom ajuste comprova que, para este caso, a escolha de um menor número de neurônios na camada oculta não traz qualquer prejuízo para a boa representação do sistema. Um fator importante que merece atenção é a qualidade dos dados, que poderia ser melhorada, caso fosse feito um acompanhamento do comportamento das variáveis no dia a dia de operação da unidade, bem como manutenção nos instrumentos de medição, principalmente nos analisadores. No caso específico do reator, seria necessário também um melhor controle da temperatura de entrada no reator, uma vez que a mesma é de fundamental importância para a reação devido a grande variação do teor de CPD (ciclopentadieno) na carga.

Os demais subsistemas serão apresentados de modo análogo, seguindo a mesma seqüência de procedimentos e apresentação dos resultados. Apesar de repetitivo, cada subsistema possui suas próprias peculiaridades e informações valiosas para o estudo.

4.5.2.2. Subsistema 2 - REDE 2

O subsistema 2 corresponde à Primeira Coluna de Remoção de Pesados (COL-01), cujo esquema de fluxo para modelagem é apresentado na Fig. 4.15, enquanto que o esquema da rede neural representativa do sistema, com as respectivas variáveis de entrada e saída, é mostrado na Fig. 4.16. As faixas de aplicabilidade destas variáveis podem ser vistas na Tabela 4.6.

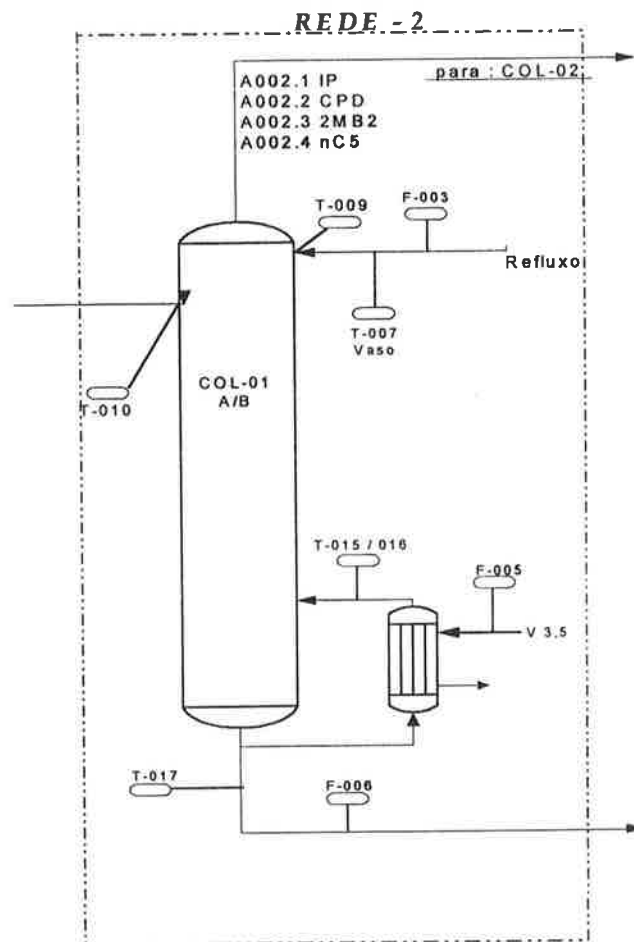


Fig. 4.15. – Fluxograma da COL-01 modelada pela REDE 2

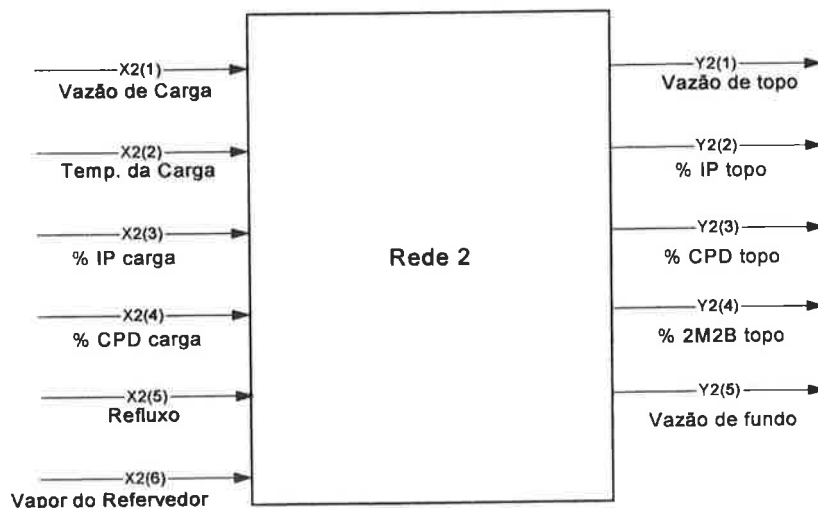


Fig. 4.16 - Esquema da REDE 2 – variáveis de entrada/saída

Tabela 4.6. - Variáveis de Processo para a REDE 2 e Suas Faixas de Aplicação

Notação	Variáveis	Unidades	Valores mínimos	Valores máximos
Entradas				
x2(1)	Vazão de carga	t/h	11,00	14,02
x2(2)	Temperatura da Carga	(°C)	62,61	66,51
x2(3)	Análise de IP na carga	%massa	14,50	20,16
x2(4)	Análise de CPD na carga	%massa	0,83	7,04
x2(5)	Vazão de Refluxo	t/h	31,22	53,23
x2(6)	Vapor para Refervedor	t/h	7,94	11,59
Saídas				
y2(1)	Vazão de topo	t/h	5,57	8,39
y2(2)	Análise de IP no topo	%massa	24,82	37,74
y2(3)	Análise de CPD no topo	%massa	0,85	3,49
y2(4)	Análise de 2M2B no topo	%massa	0,07	1,24
y2(5)	Vazão de fundo	t/h	4,56	6,40

Notas: IP – Isopreno; CPD – Ciclopentadieno; 2M2B – 2-Metil-2-Buteno

De acordo com a metodologia apresentada, uma vez definidas as variáveis de entrada e saída, foi realizada uma primeira etapa de treinamento, contendo todos os pontos no conjunto de aprendizagem, de modo a verificar a consistência dos dados e eliminação de dados anômalos como descrito anteriormente na Parte II do item 4.4.2.2. Neste caso, durante as etapas de estabelecimento de períodos em estado estacionário e análise e tratamento dos dados utilizando redes neurais, verificou-se uma grande dispersão nos dados, o que levou a uma redução substancial no número de pontos experimentais válidos para utilização na etapa de configuração e treinamento da rede neural. Assim, a qualidade dos dados foi um fator de grande relevância para a modelagem da coluna. Um conjunto de dados de melhor qualidade teria levado a resultados mais consistentes.

Em uma segunda etapa, seguindo a mesma metodologia já descrita anteriormente e utilizada no subsistema 1, os dados foram divididos em dois grupos: conjunto de aprendizagem e conjunto de testes e foram feitas diversas corridas variando o número de neurônios na camada oculta e os resultados da variação do erro residual do conjunto de teste em função do número de iterações e neurônios na camada oculta é apresentado na Fig. A.2 (ANEXO A).

O número ótimo de neurônios na camada oculta (NH) correspondente ao menor erro obtido para o conjunto de testes seria 15 e o número de iterações igual à cerca de 40000. Mas, neste caso, como no anterior, optou-se por uma configuração de rede com menor número de neurônios e também menor número de iterações. Assim, a rede utilizada apresenta 13 neurônios na camada oculta e 30000 iterações, sem perda de qualidade no desempenho da representação do sistema. Melhor ajuste poderia ser obtido mediante a utilização de um conjunto de dados de melhor qualidade.

A topologia ou configurações da rede neural selecionada é apresentada na Tabela 4.7 para a REDE 2.

Tabela 4.7. - Configurações da REDE 2

REDE 2	
Número de Pontos do conjunto de aprendizagem	56
Número de Pontos do conjunto de teste	19
Número de Variáveis de Entrada	6
Número de Variáveis de Saída	5
Número de Neurônios na Camada Oculta	13
Número de Neurônios na Camada de Saída	5
Número de Iterações	30000
Fator de Amortecimento	1,00
Parâmetro de Momento	1,00

Analogamente à REDE 1, os resultados do subsistema 2 (REDE 2) são apresentados a seguir, de modo a verificar o ajuste do modelo da rede neural.

As figuras que apresentam os resultados obtidos no ajuste para o conjunto de dados de teste, comparando os valores das variáveis de saída medidas e calculadas, bem como os testes de avaliação do desempenho da rede podem ser vistas no ANEXO B.

Observando os gráficos da Fig. B.4 (ANEXO B), verifica-se ainda a existência de pequena dispersão entre os valores calculados e experimentais, o que pode ser comprovado pelos fatores de correlação apresentados nos gráficos. Acredita-se que melhor ajuste poderia ser obtido utilizando dados de melhor qualidade.

Os resultados seguintes mostram os testes para verificação da consistência do modelo da rede neural gerada. O primeiro teste foi o de distribuição da função distância. A Fig. B.5 (ANEXO B) mostra os resultados deste teste.

A função distância fornece a medida da diferença entre o valor experimental e o calculado pela rede e quanto mais equitativa a distribuição, melhor o ajuste. Assim, observa-se que, de modo geral o ajuste é satisfatório, à exceção das

variáveis vazão de topo e análise de isopreno no topo, para as quais as distribuições não se mostram exatamente simétricas.

A outra forma de verificação da consistência e ajuste dos dados á representação da distribuição normal dos erros entre os valores experimentais e os gerados pela rede neural. A Fig. B.6 (ANEXO B) mostra estes resultados para o conjunto de dados de teste. Os resultados mostram um ajuste satisfatório, já que a distribuição dos erros apresenta um comportamento aproximadamente normal com média zero.

De modo geral, os erros avaliados pelas técnicas mostradas não apresentam uma variabilidade apreciável, apesar de, em alguns casos, ficar evidente um pequeno desvio do comportamento esperado. Dessa forma, o ajuste obtido através do modelo da REDE 2 pode ser considerado satisfatório para aplicações práticas.

4.5.2.3. Subsistema 3 - REDE 3

O subsistema 3 corresponde à Coluna Depentanizadora (COL-12), cujo esquema de fluxo para modelagem é apresentado na Fig. 4.17, enquanto que o esquema da rede neural representativa do sistema, com as respectivas variáveis de entrada e saída, é mostrado na Fig. 4.18. As faixas de aplicabilidade destas variáveis podem ser vistas na Tabela 4.8.

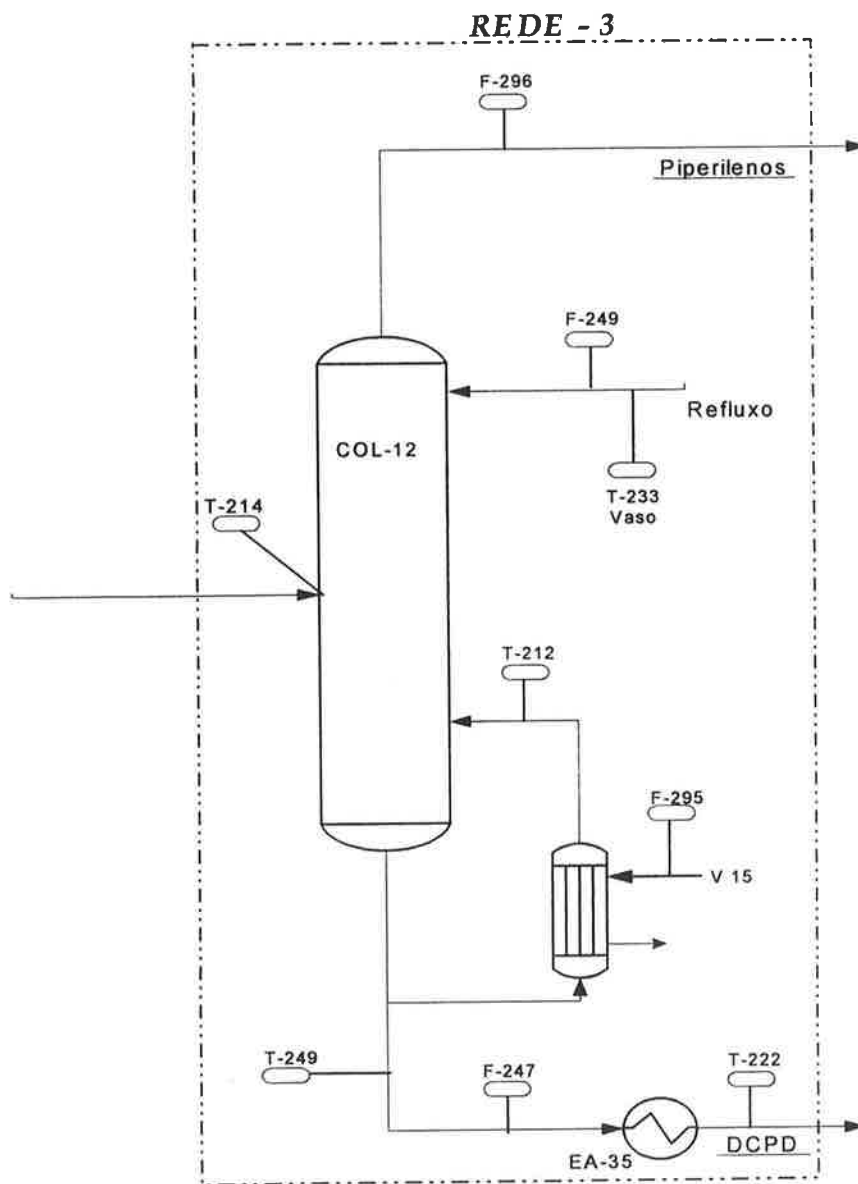


Fig. 4.17. – Fluxograma da COL-12 modelada pela REDE 3

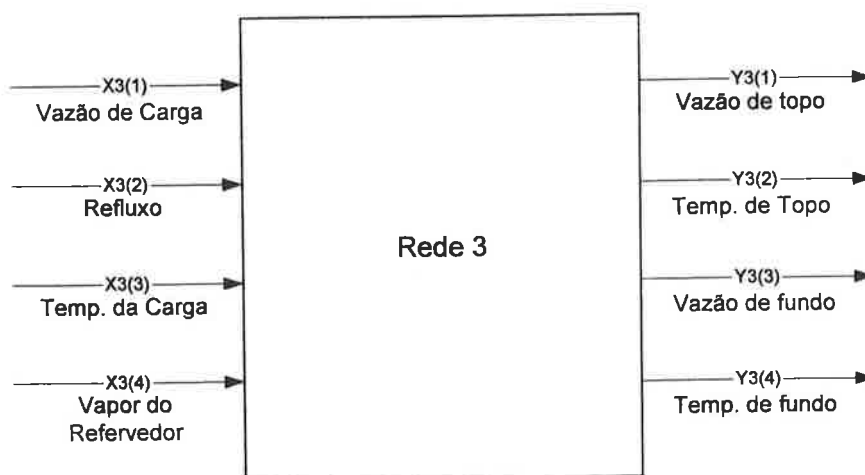


Fig. 4.18 - Esquema da REDE 3 – variáveis de entrada/saída

Tabela 4.8. Variáveis de Processo para a REDE 3 e Suas Faixas de Aplicação

Notação	Variáveis	Unidades	Valores mínimos	Valores máximos
Entradas				
x3(1)	Vazão de carga	t/h	4,61	6,11
x3(2)	Vazão de Refluxo	t/h	1,60	3,41
x3(3)	Temperatura da carga	(°C)	55,46	66,38
x3(4)	Vapor para Refervedor	t/h	0,83	1,33
Saídas				
y3(1)	Vazão de topo	t/h	2,21	3,68
y3(2)	Temperatura de topo	(°C)	48,70	51,52
y3(3)	Vazão de fundo	t/h	1,33	2,99
y3(4)	Temperatura de fundo	(°C)	110,75	126,98

Neste caso, importante observação é a não inclusão, como variáveis de saída, das análises no topo de isopreno (IP), ciclopentadieno (CPD), n-pentano (n-C5) e 2-metil-2-buteno (2M2B). Embora sejam variáveis importantes para operação da coluna, principalmente quando se trata da otimização da seção de preparação de

carga¹, não foi possível o ajuste da REDE 3 utilizando estas variáveis. As análises de IP, CPD e n-C5 apresentam valores praticamente constantes, enquanto que a análise de 2M2B apresenta um conjunto de dados com alta variabilidade, mas sem qualquer tendência. Os valores dos erros para rede incluindo estas variáveis foram bastante altos e muitas vezes apresentavam um comportamento oscilatório, independente do número de neurônios empregado. Uma vez excluído este conjunto de informações, o erro residual reduziu muito, sendo, então possível o ajuste da rede.

Uma vez definidas as variáveis de entrada e saída e analisados os dados quanto à detecção de “outliers”, seguindo a mesma metodologia já descrita anteriormente e utilizada nos subsistema 1 e 2, os dados foram divididos em dois grupos: conjunto de aprendizagem e conjunto de testes e foram feitas diversas corridas variando o número de neurônios na camada oculta.

O número ótimo de neurônios na camada oculta (NH) correspondente ao menor erro obtido para o conjunto de testes é 16 e o número de iterações igual à cerca de 45000. Os resultados da variação do erro residual do conjunto de teste em função do número de iterações e neurônios na camada oculta são apresentados na Fig. A.3 (ANEXO A). A topologia ou configurações da rede neural selecionada é apresentada na Tabela 4.9 para a REDE 3.

Tabela 4.9. - Configurações da REDE 3

REDE 3	
Número de Pontos do conjunto de aprendizagem	97
Número de Pontos do conjunto de teste	37
Número de Variáveis de Entrada	4
Número de Variáveis de Saída	4
Número de Neurônios na Camada Oculta	16
Número de Neurônios na Camada de Saída	4
Número de Iterações	45000
Fator de Amortecimento	1,00
Parâmetro de Momento	1,00

¹ A seção de preparação de carga é constituída pelo Reator de Dimerização (REDE 1), Primeira Coluna de Remoção de Pesados (REDE 2) e Coluna Depentanizadora (REDE 3).

Analogamente aos subsistemas 1 e 2, os resultados do subsistema 3 (REDE 3) são apresentados a seguir, de modo a verificar o ajuste do modelo da rede neural. As figuras que apresentam os resultados obtidos no ajuste para o conjunto de dados de teste, comparando os valores das variáveis de saída medidas e calculadas, bem como os testes de avaliação da consistência e desempenho do modelo da rede gerada podem ser vistas no ANEXO B.

De maneira geral, os ajustes podem ser considerados satisfatórios apesar da pequena dispersão apresentada principalmente nas temperaturas, como pode ser visto pelos fatores de correlação ambos abaixo de 0,9. Deve-se ter em mente que, apesar de ter sido realizado um tratamento rigoroso nos dados, estão sendo utilizados dados operacionais reais.

A função distância fornece a medida da diferença entre o valor experimental e o calculado pela rede e quanto mais equitativa a distribuição, melhor o ajuste. Assim, observa-se que, de modo geral o ajuste é satisfatório, à exceção da variável vazão de fundo, para a qual a distribuição encontra-se deslocada em direção ao sentido positivo dos erros.

A outra forma de verificação da consistência e ajuste dos dados é representação da distribuição normal dos erros entre os valores experimentais e os gerados pela rede neural, a qual indica, para este caso, mostram um bom ajuste. já que a distribuição dos erros para todas as variáveis apresenta um comportamento aproximadamente normal com média zero.

De modo geral, os erros avaliados pelas técnicas mostradas apresentam evidenciam um bom ajuste das variáveis modeladas através da REDE 3 proposta, respeitando as restrições observadas para este caso.

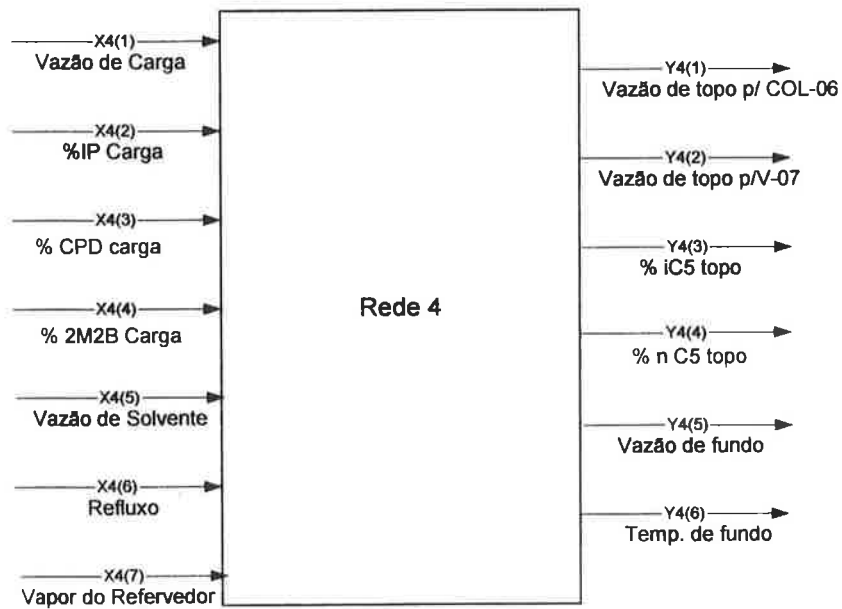


Fig. 4.20 - Esquema da REDE 4 – variáveis de entrada/saída

Tabela 4.10. - Variáveis de Processo para a REDE 4 e Suas Faixas de Aplicação

Notação	Variáveis	Unidades	Valores mínimos	Valores máximos
Entradas				
x4(1)	Vazão de carga	t/h	5,79	8,54
x4(2)	Análise de IP na carga	%massa	21,30	38,54
x4(3)	Análise de CPD na carga	%massa	0,81	3,66
x4(4)	Análise de 2M2B na carga	%massa	0,03	1,38
x4(5)	Vazão de Solvente	t/h	53,37	71,49
x4(6)	Vazão de Refluxo	t/h	10,62	17,67
x4(7)	Vapor para Refervedor	t/h	2,48	5,68
Saídas				
y4(1)	Vazão de topo I	t/h	1,61	3,70
y4(2)	Vazão de topo II	t/h	3,15	5,13
y4(3)	Análise de i-C5 no topo	%massa	19,26	39,73
y4(4)	Análise de n-C5 no topo	%massa	22,48	39,61
y4(5)	Vazão de fundo	t/h	51,43	71,10
y4(6)	Temperatura de fundo	(°C)	84,95	91,23

Neste caso, assim como para o subsistema 3, importante observação é a não inclusão, como variáveis de saída, das análises no topo de isopreno (IP), pentadieno-1,4 (PTD1,4), e 2-metil-2-buteno (2M2B). Embora sejam variáveis importantes para a operação da coluna, principalmente quando se trata da sua otimização em relação à presença de compostos que poderiam contaminar o produto final ou à perda de IP, não desejável, não foi possível o ajuste da REDE 4 utilizando estas variáveis. As análises destes compostos apresentaram valores muito baixos e praticamente constantes. Os valores dos erros para a rede incluindo estas variáveis foram bastante altos e, muitas vezes, apresentavam um comportamento oscilatório, independente do número de neurônios empregado. Uma vez excluído este conjunto de informações, o erro residual foi reduzido substancialmente, sendo, então possível o ajuste da rede.

Uma vez definidas as variáveis de entrada e saída e analisados os dados, seguindo a mesma metodologia já descrita anteriormente, os dados foram divididos em dois grupos: conjunto de aprendizagem e conjunto de testes e foram feitas diversas corridas variando o número de neurônios na camada oculta. O número ótimo de neurônios na camada oculta (NH) correspondente ao menor erro obtido para o conjunto de testes é 19 e o número de iterações igual à cerca de 30000 é satisfatório.

Os resultados da variação do erro residual do conjunto de teste em função do número de iterações e neurônios na camada oculta são apresentados no ANEXO A.

A topologia ou configurações da rede neural selecionada é apresentada na Tabela 4.11 para a REDE 4.

Tabela 4.11. - Configurações da REDE 4

REDE 4	
Número de Pontos do Conjunto de Aprendizagem	109
Número de Pontos do Conjunto de Teste	42
Número de Variáveis de Entrada	7
Número de Variáveis de Saída	6
Número de Neurônios na Camada Oculta	19
Número de Neurônios na Camada de Saída	6
Número de Iterações	30000
Fator de Amortecimento	1,00
Parâmetro de Momento	1,00

Analogamente aos casos anteriores, os resultados do subsistema 4 (REDE 4) são discutidos a seguir, de modo a verificar o ajuste do modelo. As figuras que apresentam os resultados obtidos no ajuste para o conjunto de dados de teste, comparando os valores das variáveis de saída medidas e calculadas, bem como os testes de avaliação da consistência e desempenho do modelo da rede gerada podem ser vistas no ANEXO B.

De maneira geral, os ajustes podem ser considerados satisfatórios apesar da pequena dispersão apresentada principalmente nas Temperaturas, como pode ser visto pelos fatores de correlação ambos abaixo de 0,9. Deve-se ter em mente que, apesar de ter sido realizado um tratamento rigoroso nos dados, estão sendo utilizados dados operacionais reais.

Os resultados seguintes mostram os testes para verificação da consistência do modelo da rede neural gerada. O primeiro teste foi o de distribuição da função distância, a qual fornece a medida da diferença entre o valor experimental e o calculado pela rede. Quanto mais eqüitativa a distribuição, melhor o ajuste. Observa-se que apesar da ligeira dispersão verificada para algumas variáveis de saída como visto na Fig. B.10 (ANEXO B), o teste da função distância apresenta distribuições praticamente simétricas, à exceção da variável temperatura de

fundo, para a qual a distribuição encontra-se deslocada em direção ao sentido positivo dos erros. Assim, de modo geral, o ajuste da rede pode ser considerado bom.

A outra forma de verificação da consistência e ajuste dos dados é a representação da distribuição normal dos erros entre os valores experimentais e os gerados pela rede neural.

A partir deste critério, pode-se observar que todas as variáveis apresentam ajustes satisfatórios, já que a distribuição dos erros é aproximadamente normal.

O comportamento geral das variáveis é bem representada pelo modelo da rede neural proposta, evidenciando um ajuste satisfatório dentro das restrições para este caso. Assim, levando em consideração o ajuste e seus testes de avaliação de consistência, as variáveis de saída foram bem representadas pela REDE 4.

4.5.2.5. Subsistema 5 - REDE 5

O subsistema 5 corresponde à Segunda Coluna de Destilação Extrativa (COL-03) e à Coluna Stripper de Solvente (COL-04), cujo esquema de fluxo para modelagem é apresentado na Fig. 4.21, enquanto que o esquema da rede neural representativa do sistema, com as respectivas variáveis de entrada e saída, é mostrado na Fig. 4.22. As faixas de aplicabilidade destas variáveis podem ser vistas na Tabela 4.12. A REDE 5 modela as duas colunas referenciadas com um único sistema.

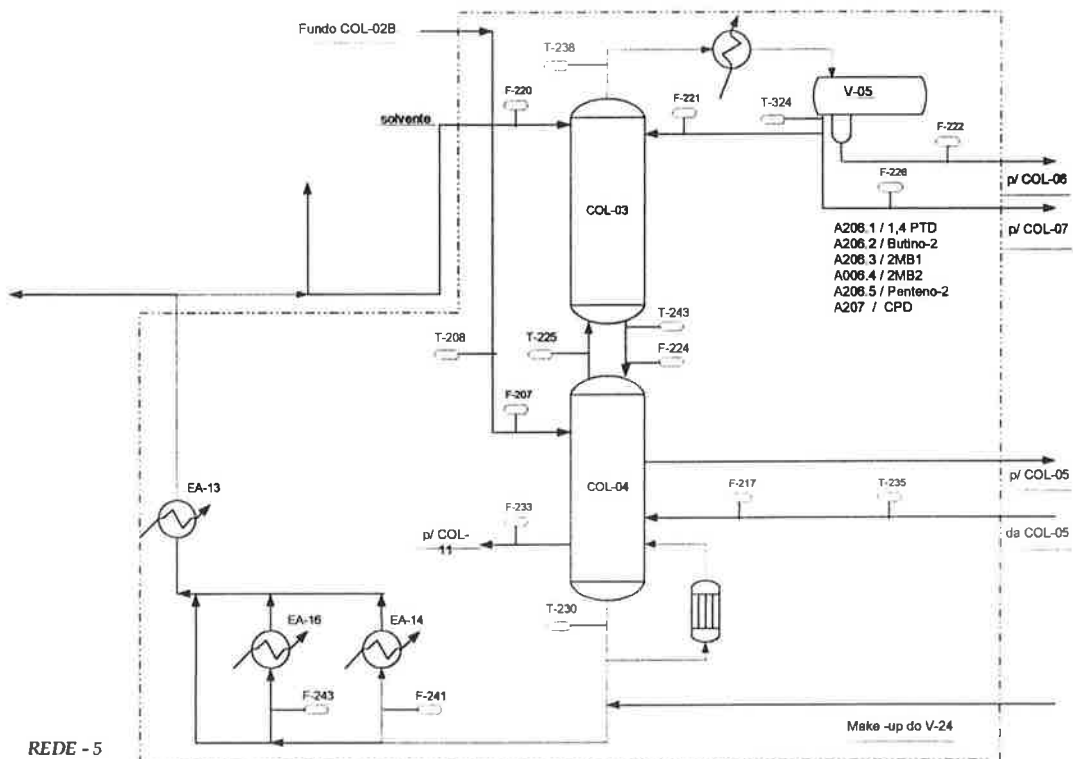


Fig. 4.21. – Fluxograma das COL-03 e COL-04 modeladas pela REDE 5

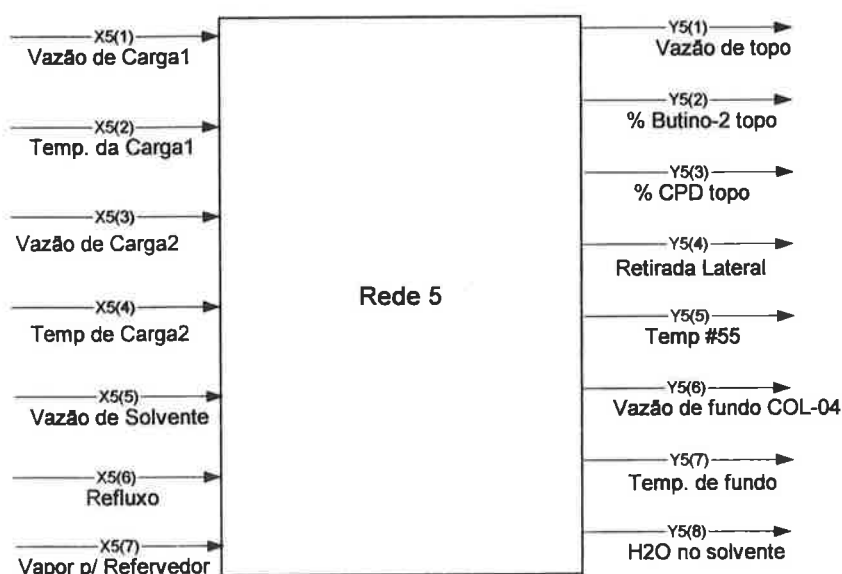


Fig. 4.22 - Esquema da REDE 5 – variáveis de entrada/saída

Tabela 4.12. Variáveis de Processo para a REDE 5 e Suas Faixas de Aplicação

Notação	Variáveis	Unidades	Valores mínimos	Valores máximos
Entradas				
x5(1)	Vazão de Carga I	t/h	57,27	68,63
x5(2)	Temperatura de Carga I	(°C)	85,10	91,23
x5(3)	Vazão de Carga II	t/h	0,77	1,33
x5(4)	Temperatura de Carga II	(°C)	90,78	95,14
x5(5)	Vazão de Solvente	t/h	13,00	16,90
x5(6)	Vazão de Refluxo	t/h	12,40	15,91
x5(7)	Vapor para Refervedor	t/h	4,82	6,20
Saídas				
y5(1)	Vazão de Topo	t/h	2,20	3,00
y5(2)	Análise de Butino-2 no topo	%massa	0,88	1,68
y5(3)	Análise de CPD no topo	ppm	1,20	5,94
y5(4)	Retirada Lateral	t/h	1,22	1,89
y5(5)	Temperatura # 55	(°C)	98,22	101,07
y5(6)	Vazão de Fundo	t/h	68,77	80,15
y5(7)	Temperatura de Fundo	(°C)	101,90	104,06
y5(8)	Análise de H2O no Solvente	%massa	15,09	26,17

Durante o tratamento dos dados para esta rede, também foram verificados problemas de convergência, quando da inclusão das seguintes variáveis de saída: análises de pentadieno-1,4 (PTD1,4), 2-metil-1-buteno (2M1B) e 2-metil-2-buteno (2M2B). Embora sejam variáveis importantes para a operação da coluna, principalmente quando se trata da sua otimização em relação à presença de compostos que poderiam contaminar o produto final, não foi possível o ajuste da REDE 5 utilizando estas variáveis. As análises destes compostos apresentaram valores muito baixos e praticamente constantes. Os valores dos erros para a rede incluindo estas variáveis foram bastante altos e, muitas vezes, apresentavam um comportamento oscilatório, independente do número de neurônios empregado. Uma vez excluído este conjunto de informações, o erro residual foi reduzido substancialmente, sendo, então possível o ajuste da rede.

Uma vez definidas as variáveis de entrada e saída e analisados os dados, seguindo a mesma metodologia já descrita anteriormente, os dados foram divididos em dois grupos: conjunto de aprendizagem e conjunto de testes e foram feitas diversas corridas variando o número de neurônios na camada oculta e os gráficos mostrando os resultados da variação do erro residual do conjunto de teste em função do número de iterações e neurônios na camada oculta são apresentados no ANEXO A.

O número ótimo de neurônios na camada oculta (NH), correspondente ao menor erro obtido para o conjunto de testes, é igual a 15. No entanto optou-se por uma rede com 14 neurônios, pois, como pode ser visto, não há diferença apreciável entre os erros residuais para as duas configurações de rede. O número de iterações igual a 67000 é satisfatório, apresentando um melhor ajuste do sistema.

A topologia ou configurações da rede neural selecionada é apresentada na Tabela 4.13 para a REDE 5.

Tabela 4.13. - Configurações da REDE 5

REDE 5	
Número de Pontos do Conjunto de Aprendizagem	79
Número de Pontos do Conjunto de Teste	27
Número de Variáveis de Entrada	7
Número de Variáveis de Saída	8
Número de Neurônios na Camada Oculta	14
Número de Neurônios na Camada de Saída	8
Número de Iterações	67000
Fator de Amortecimento	0,85
Parâmetro de Momento	1,00

Analogamente aos casos anteriores, os resultados do subsistema 5 (REDE 5) são apresentados a seguir, de modo a verificar o ajuste do modelo da rede neural. As figuras que apresentam os resultados obtidos no ajuste para o conjunto de dados de teste, comparando os valores das variáveis de saída medidas e calculadas, bem como os testes de avaliação da consistência e desempenho do modelo da rede gerada podem ser vistas no ANEXO B.

Observa-se que os ajustes das variáveis do topo da COL-03 (vazão e composição) não apresentam um bom ajuste, visto a grande dispersão dos dados, o que é comprovado pelos baixos fatores de correlação, embora tenha sido realizado um trabalho rigoroso de tratamento dos dados operacionais. No entanto, as variáveis relacionadas à COL-04, de maneira geral, apresentam um bom ajuste.

Os resultados seguintes mostram os testes para verificação da consistência do modelo da rede neural gerada. O primeiro teste foi o de distribuição da função distância.

Os gráficos da Fig. B.14 (ANEXO B), representando a função distância apresentam descolamentos do ponto central, indicativo de uma distribuição não equitativa. Estes resultados vêm ao encontro daqueles já visualizados nos

gráficos da Fig. B.13, comparando os dados experimentais e os valores calculados pela rede. Fica, então, evidenciado que o ajuste poderia ser melhorado. No entanto, tratamento nos dados e seleção das variáveis da rede foram etapas realizadas de forma rigorosa e criteriosa, não sendo, portanto, possível obter melhores resultados com o conjunto de dados utilizados. Deve-se ter em mente que se trata de dados operacionais reais.

A outra forma de verificação da consistência e ajuste dos dados é a representação da distribuição normal dos erros entre os valores experimentais e os gerados pela rede neural. As figuras dos histogramas da distribuição dos erros para o conjunto de dados de teste são encontradas no ANEXO B.

A partir deste critério, pode-se observar que todas as variáveis apresentam ajustes satisfatórios, já que a distribuição dos erros é aproximadamente normal.

O comportamento geral das variáveis é bem representada pelo modelo da rede neural proposta, evidenciando um ajuste satisfatório dentro das restrições para este caso. Assim, levando em consideração o ajuste e seus testes de avaliação de consistência, as variáveis de saída foram satisfatoriamente representadas pela REDE 5.

4.5.2.6. Subsistema 6 – REDE 6

O subsistema 6 corresponde à Coluna de Remoção de Ciclopentadieno (CPD) (COL-05), cujo esquema de fluxo para modelagem é apresentado na Fig. 4.23, enquanto que o esquema da rede neural representativa do sistema, com as respectivas variáveis de entrada e saída, é mostrado na Fig. 4.24. As faixas de aplicabilidade destas variáveis podem ser vistas na Tabela 4.14.

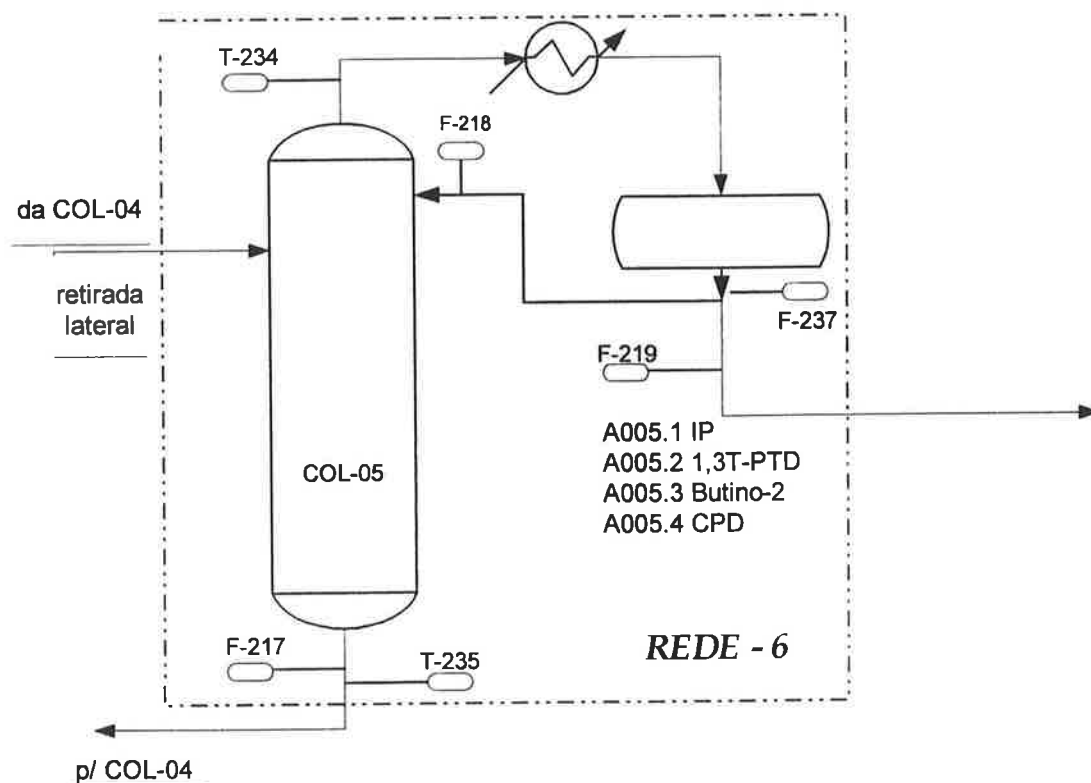


Fig. 4.23. – Fluxograma da COL-05 modelada pela REDE 6

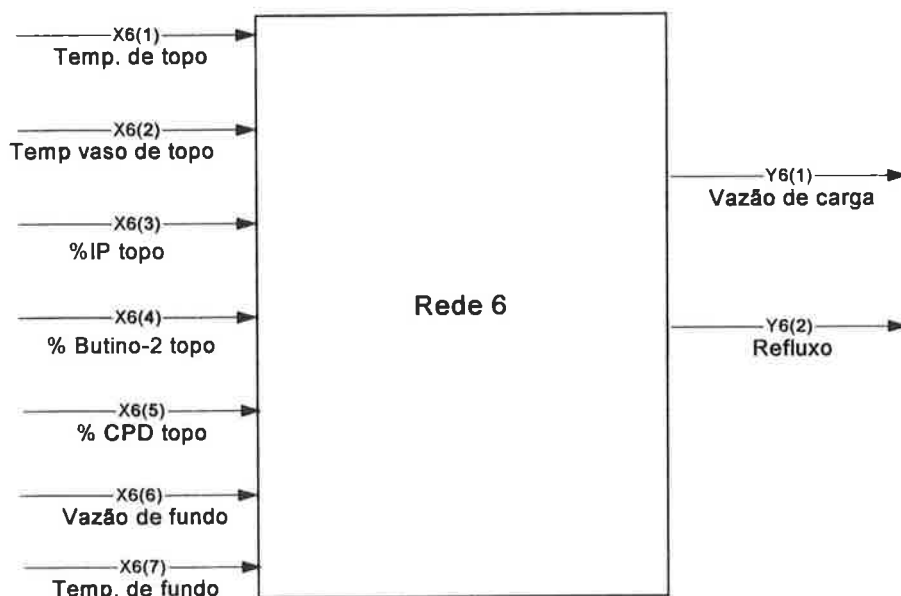


Fig. 4.24 - Esquema da REDE 6 – variáveis de entrada/saída

Tabela 4.14. - Variáveis de Processo para a REDE 6 e Suas Faixas de Aplicação

Notação	Variáveis	Unidades	Valores mínimos	Valores máximos
Entradas				
x6(1)	Temperatura de Topo	(°C)	54,81	76,96
x6(2)	Temperatura do Vaso de Topo	(°C)	27,45	32,03
x6(3)	Análise de IP no Topo	%massa	1,56	18,71
x6(4)	Análise de Butino-2 no Topo	%massa	10,16	19,87
x6(5)	Análise de CPD no Topo	%massa	20,84	58,91
x6(6)	Vazão de Fundo	t/h	0,79	1,37
x6(7)	Temperatura de Fundo	(°C)	89,4	95,75
Saídas				
y6(1)	Vazão de Carga	t/h	1,24	2,00
y6(2)	Vazão de Refluxo	t/h	1,17	1,84

Neste caso, para a modelagem da COL-05, observa-se que as variáveis de entrada e saída da rede encontram-se invertidas em relação ao fluxo de processamento apresentado na Fig. 4.23. Isto foi devido à impossibilidade de modelagem seguindo o fluxo convencional. Observa-se, na Tabela 4.14, que a variação das vazões de carga e refluxo (variáveis de entrada reais) são pequenas e não se conseguiu discriminar as possíveis variáveis de saída (variáveis reais de saída) apenas com estas informações, obtendo-se valores altos para o erro residual no ajuste das redes propostas, independente do número de neurônios empregado. Uma vez invertendo as informações de entrada e saída, o erro residual foi reduzido substancialmente, sendo, então possível o ajuste da rede. Assim, a REDE 6 representa a COL-05 de maneira invertida.

Uma vez definidas as variáveis de entrada e saída e analisados os dados, seguindo a mesma metodologia já descrita anteriormente, os dados foram divididos em dois grupos: conjunto de aprendizagem e conjunto de testes e foram feitas diversas corridas variando o número de neurônios na camada oculta e os resultados da variação do erro residual do conjunto de teste em função do número de iterações e neurônios na camada oculta são apresentados no ANEXO A.

O número ótimo de neurônios na camada oculta (NH) correspondente ao menor erro obtido para o conjunto de testes pode ser 12, 13 ou 15 e número de iterações acima de 50000 . Assim, optou-se por uma rede com 12 neurônios, pois, como pode ser visto, não há diferença apreciável entre os erros residuais para estas configurações de rede. O número de iterações igual a 60000 é empregado de modo a assegurar um melhor ajuste do sistema.

A topologia ou configurações da rede neural selecionada é apresentada na Tabela 4.15 para a REDE 6.

Tabela 4.15. - Configurações da REDE 6

REDE 6	
Número de Pontos do Conjunto de Aprendizagem	225
Número de Pontos do Conjunto de Teste	106
Número de Variáveis de Entrada	7
Número de Variáveis de Saída	2
Número de Neurônios na Camada Oculta	12
Número de Neurônios na Camada de Saída	2
Número de Iterações	60000
Fator de Amortecimento	1,00
Parâmetro de Momento	1,00

Analogamente aos casos anteriores, os resultados do subsistema 6 (REDE 6) são apresentados a seguir, de modo a verificar o ajuste do modelo da rede neural. As figuras que apresentam os resultados obtidos no ajuste para o conjunto de dados de teste, comparando os valores das variáveis de saída medidas e calculadas, bem como os testes de avaliação da consistência e desempenho do modelo da rede gerada podem ser vistas no ANEXO B.

Boa Concordância entre os valores modelados e medidos podem ser observados e a distribuição uniforme indica ausência de qualquer tendência no ajuste.

Os resultados seguintes mostram os testes para verificação da consistência do modelo da rede neural gerada. O primeiro teste foi o de distribuição da função distância. Pode-se perceber, pela distribuição equitativa, que o modelo apresenta um bom ajuste para ambas as variáveis de saída.

Outra forma de apresentação é a distribuição normal dos erros entre os valores calculados pela rede neural e os experimentais. A distribuição normal demonstra o equilíbrio de frequência dos valores encontrados em torno da média zero. Uma distribuição normal mostra a boa qualidade do ajuste realizado. Os histogramas da distribuição de erros para as variáveis de saída da REDE 6 para o

conjunto de dados de teste mostram que o ajuste é muito bom, já que a distribuição dos erros apresenta um comportamento normal com média aproximadamente igual a zero.

De modo geral, os erros avaliados pelas técnicas mostradas não apresentam uma variabilidade apreciável, evidenciando, portanto, um bom ajuste do modelo da REDE 6.

4.5.2.7. Subsistema 7 - REDE 7

O subsistema 7 corresponde à Coluna Lavadora de C5 (COL-06) e à Coluna Lavadora de Produto (COL-07), cujo esquema de fluxo para modelagem é apresentado na Fig. 4.25, enquanto que o esquema da rede neural representativa do sistema, com as respectivas variáveis de entrada e saída, é mostrado na Fig. 4.26. A REDE 7 modela as colunas COL-06 e COL-07 como um único sistema. As faixas de aplicabilidade destas variáveis podem ser vistas na Tabela 4.16.

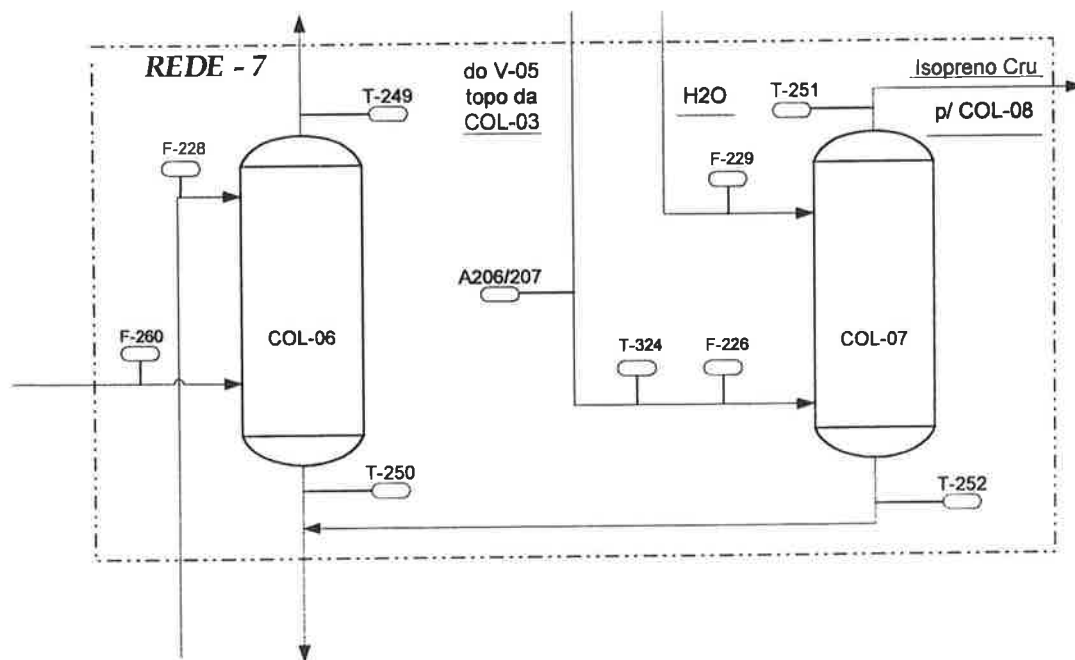


Fig. 4.25. – Fluxograma das COL-06 e COL-07 modeladas pela REDE 7

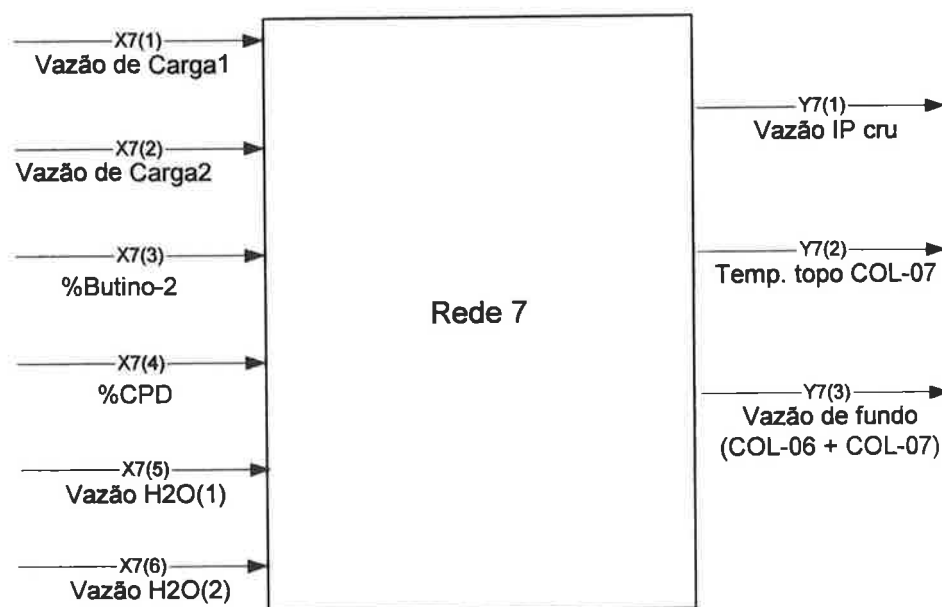


Fig. 4.26. - Esquema da REDE 7 – variáveis de entrada/saída

Tabela 4.16. - Variáveis de Processo para a REDE 7 e Suas Faixas de Aplicação

Notação	Variáveis	Unidades	Valores mínimos	Valores máximos
Entradas				
x7(1)	Vazão de Carga I	t/h	4,45	8,71
x7(2)	Vazão de Carga II	t/h	2,00	3,15
x7(3)	Análise de Butino-2 na Carga	%massa	0,34	2,18
x7(4)	Análise de CPD na Carga	ppm	0,50	6,61
x7(5)	Vazão de Água I	t/h	1,90	3,08
x7(6)	Vazão de Água II	t/h	1,30	3,05
Saídas				
y7(1)	Vazão de Topo (IP cru)	t/h	1,68	2,83
y7(2)	Temperatura de Topo	(°C)	26,80	31,60
y7(3)	Vazão de Fundo	t/h	4,20	7,30

Uma vez definidas as variáveis de entrada e saída e analisados os dados, seguindo a mesma metodologia já descrita anteriormente, os dados foram divididos em dois grupos: conjunto de aprendizagem e conjunto de testes e foram feitas diversas corridas variando o número de neurônios na camada oculta e os resultados da variação do erro residual do conjunto de teste em função do número de iterações e neurônios na camada oculta são apresentados no ANEXO A.

O número ótimo de neurônios na camada oculta (NH) correspondente ao menor erro obtido para o conjunto de testes pode ser 13 ou 16. Optou-se por uma rede com 13 neurônios, pois, como pode ser visto, não há diferença apreciável entre os erros residuais para estas configurações de rede. O número de iterações empregado corresponde ao erro mínimo e é igual a 59500.

A topologia ou configurações da rede neural selecionada é apresentada na Tabela 4.17 para a REDE 7.

Tabela 4.17. - Configurações da REDE 7

REDE 7	
Número de Pontos do Conjunto de Aprendizagem	282
Número de Pontos do Conjunto de Teste	136
Número de Variáveis de Entrada	6
Número de Variáveis de Saída	3
Número de Neurônios na Camada Oculta	13
Número de Neurônios na Camada de Saída	3
Número de Iterações	59500
Fator de Amortecimento	1,00
Parâmetro de Momento	1,00

Analogamente aos casos anteriores, os resultados do subsistema 7 (REDE 7) são apresentados a seguir, de modo a verificar o ajuste do modelo da rede neural. As figuras que apresentam os resultados obtidos no ajuste para o conjunto de dados de teste, comparando os valores das variáveis de saída medidas e

calculadas, bem como os testes de avaliação da consistência e desempenho do modelo da rede gerada podem ser vistas no ANEXO B.

Observa-se que o ajuste da vazão de topo da COL-07 é bastante satisfatório, apresentando um bom ajuste de correlação. O ajuste da vazão de fundo das duas colunas combinadas apesar de apresentar um bom ajuste, sem grandes dispersões, encontra-se deslocado em relação à linha de 45°C. O ajuste da temperatura de topo da COL-07 não pode ser considerado bom, porém foram os melhores resultados conseguidos. Observa-se que neste caso a faixa de variação da medida é muito estreita, o que dificulta o ajuste, no entanto a eliminação desta variável leva a piores resultados, inclusive no ajuste das demais variáveis. Ao mesmo tempo, os pontos que se encontram dispersos nos demais casos não puderam ser eliminados de acordo com o tratamento de dados realizados, conforme descrito anteriormente. Assim, este foi o melhor ajuste conseguido para este conjunto de dados.

Os resultados seguintes mostram os testes para verificação da consistência do modelo da rede neural gerada. O primeiro teste foi o de distribuição da função distância.

Observando-se os gráficos da função distância para estas variáveis, pode-se chegar às mesmas conclusões obtidas através da análise dos gráficos de dispersão (Fig. B.19 – ANEXO B).

A outra forma de verificação da consistência e ajuste dos dados é a representação da distribuição normal dos erros entre os valores experimentais e os gerados pela rede neural. Pela análise dos histogramas, verifica-se distribuição normal para o erro de todas as variáveis de saída para a REDE 7, embora apenas o erro da vazão de topo apresente média aproximadamente igual a zero.

4.5.2.8. Subsistema 8 - REDE 8

O subsistema 8 corresponde à Primeira Coluna de Recuperação de Solvente (COL-10), cujo esquema de fluxo para modelagem é apresentado na Fig. 4.27, enquanto que o esquema da rede neural representativa do sistema, REDE 8, com as respectivas variáveis de entrada e saída, é mostrado na Fig. 4.28. As faixas de aplicabilidade destas variáveis podem ser vistas na Tabela 4.18.

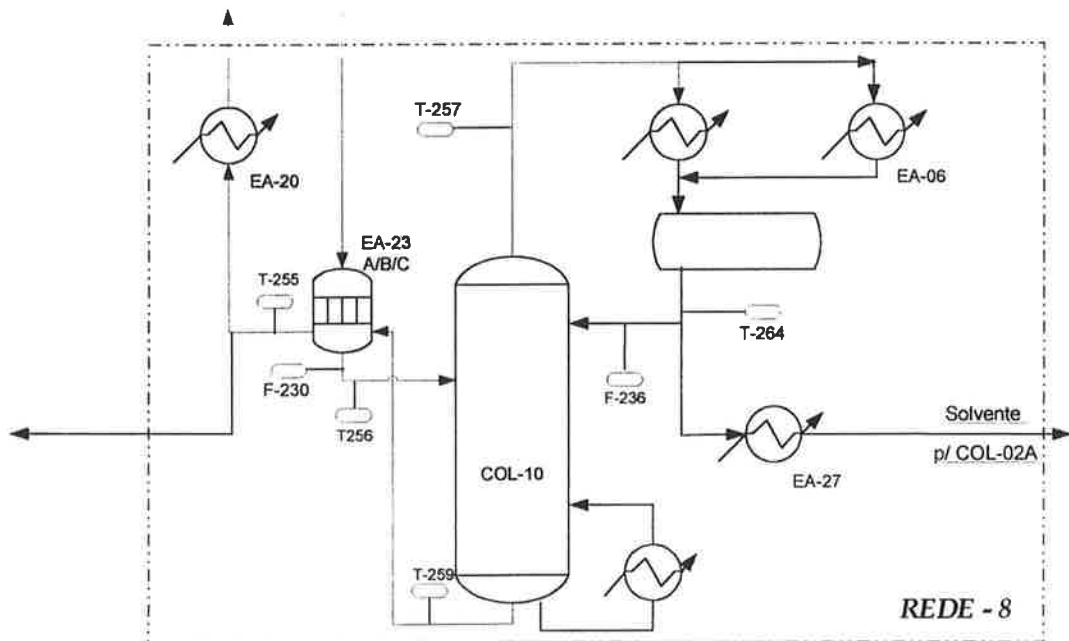


Fig. 4.27. – Fluxograma da COL-10 modelada pela REDE 8

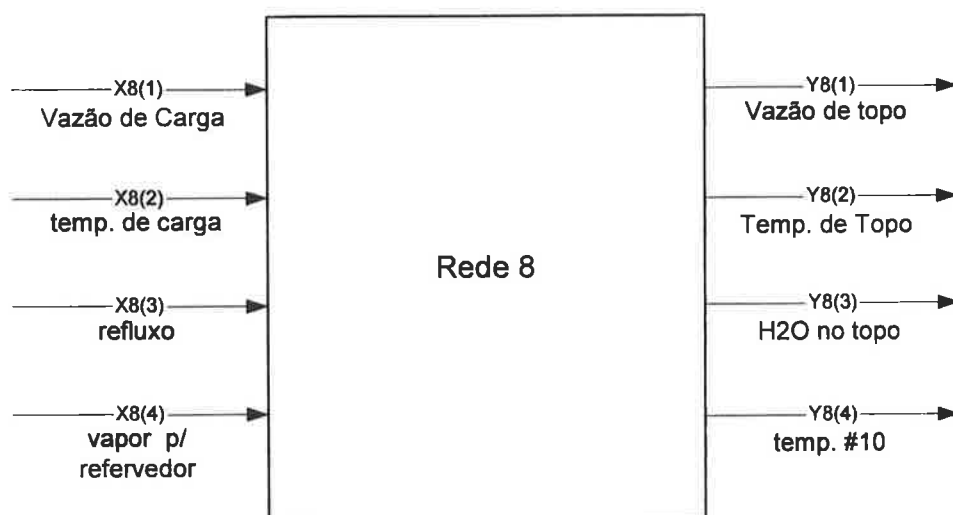


Fig. 4.28 - Esquema da REDE 8 – variáveis de entrada/saída

Tabela 4.18. - Variáveis de Processo para a REDE 8 e Suas Faixas de Aplicação

Notação	Variáveis	Unidades	Valores mínimos	Valores máximos
Entradas				
x8(1)	Vazão de Carga	t/h	3,31	7,34
x8(2)	Temperatura de Carga	(°C)	76,78	87,16
x8(3)	Vazão de Refluxo	t/h	0,60	1,95
x8(4)	Vapor para Refervedor	t/h	1,26	1,77
Saídas				
y8(1)	Vazão de topo	t/h	1,17	1,79
y8(2)	Temperatura de topo	(°C)	101,27	106,43
y8(3)	Análise de H2O no topo	% massa	21,33	30,13
y8(4)	Temperatura # 10	(°C)	105,96	110,21

Uma vez definidas as variáveis de entrada e saída e analisados os dados, seguindo a mesma metodologia já descrita anteriormente, os dados foram divididos em dois grupos: conjunto de aprendizagem e conjunto de testes e foram feitas diversas corridas variando o número de neurônios na camada oculta e os

resultados da variação do erro residual do conjunto de teste em função do número de iterações e neurônios na camada oculta são apresentados no ANEXO A.

O número ótimo de neurônios na camada oculta (NH) correspondente ao menor erro obtido para o conjunto de testes é igual a 16. O número de iterações empregado é igual a 55000, uma vez que acima deste valor o ganho na diminuição do erro não é significativo o suficiente para justificar a utilização de um maior número de iterações.

A topologia ou configurações da rede neural selecionada é apresentada na Tabela 4.19 para a REDE 8.

Tabela 4.19. - Configurações da REDE 8

REDE 8	
Número de Pontos do Conjunto de Aprendizagem	202
Número de Pontos do Conjunto de Teste	95
Número de Variáveis de Entrada	4
Número de Variáveis de Saída	4
Número de Neurônios na Camada Oculta	16
Número de Neurônios na Camada de Saída	4
Número de Iterações	55000
Fator de Amortecimento	1,00
Parâmetro de Momento	1,00

Analogamente aos casos anteriores, os resultados do subsistema 8 (REDE 8) são apresentados a seguir, de modo a verificar o ajuste do modelo da rede neural. As figuras que apresentam os resultados obtidos no ajuste para o conjunto de dados de teste, comparando os valores das variáveis de saída medidas e calculadas, bem como os testes de avaliação da consistência e desempenho do modelo da rede gerada podem ser vistas no ANEXO B.

De maneira geral, os ajustes podem ser considerados satisfatórios apesar da pequena dispersão apresentada. Deve-se ter em mente que, apesar de ter sido realizado um tratamento rigoroso nos dados, estão sendo utilizados dados operacionais reais.

Os resultados seguintes mostram os testes para verificação da consistência do modelo da rede neural gerada. O primeiro teste foi o de distribuição da função distância. Pode-se perceber, pela distribuição eqüitativa, que o modelo apresenta um bom ajuste para todas as variáveis de saída.

Outra forma de apresentação é a Distribuição Normal dos erros entre os valores calculados pela rede neural e os experimentais. Pela análise dos histogramas, verifica-se distribuição normal e média aproximadamente igual a zero para o erro residual de todas as variáveis de saída para a REDE 8, evidenciando um bom ajuste.

4.5.2.9. Subsistema 9 - REDE 9

O subsistema 9 corresponde à Coluna de Remoção de Leves (COL-08), cujo esquema de fluxo para modelagem é apresentado na Fig. 4.29, enquanto que o esquema da rede neural representativa do sistema, REDE 9, com as respectivas variáveis de entrada e saída, é mostrado na Fig. 4.30. As faixas de aplicabilidade destas variáveis podem ser vistas na Tabela 4.20.

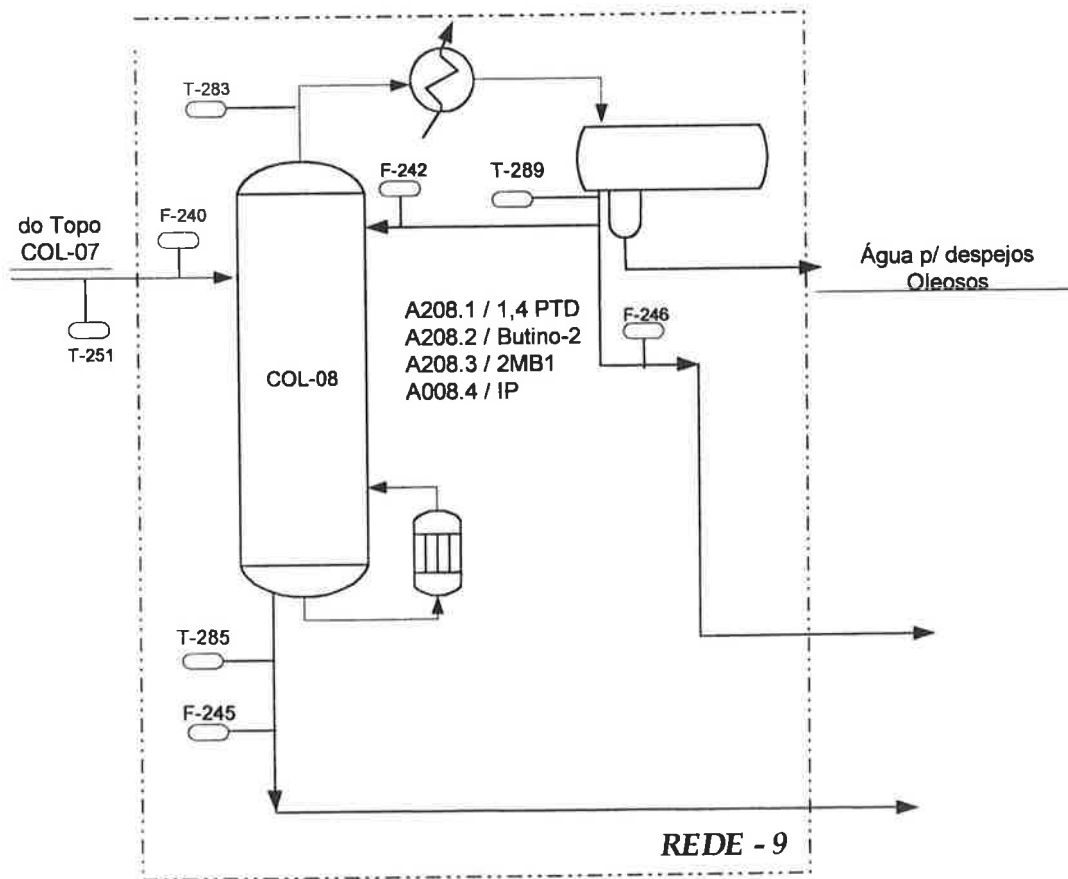


Fig. 4.29 - Fluxograma da COL-08 modelada pela REDE 9

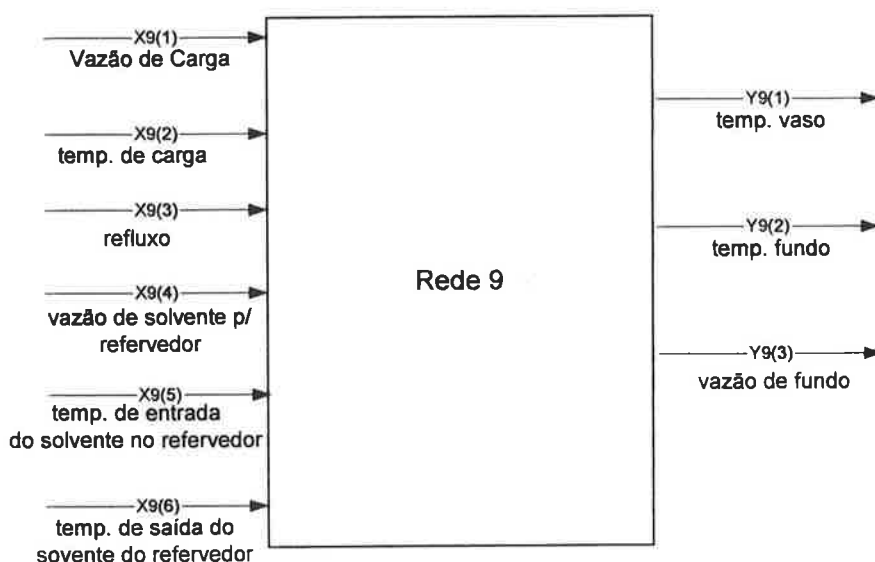


Fig. 4.30 - Esquema da REDE 9– variáveis de entrada/saída

Tabela 4.20. - Variáveis de Processo para a REDE 9 e Suas Faixas de Aplicação

Notação	Variável	Unidades	Valores mínimos	Valores máximos
Entradas				
x9(1)	Vazão de carga	t/h	1,47	3,06
x9(2)	Temperatura de Carga	(°C)	27,58	35,52
x9(3)	Vazão de Refluxo	t/h	9,36	13,88
x9(4)	Vazão de Solvente p/ Refervedor	t/h	38,26	63,98
x9(5)	Temperatura de Entrada do Solvente no Refervedor	(°C)	102,05	104,30
x9(6)	Temperatura de Saída do Solvente do Refervedor	(°C)	66,08	79,63
Saídas				
y9(1)	Temperatura do vaso	(°C)	31,73	42,73
y9(2)	Temperatura de Fundo	(°C)	58,72	64,06
y9(3)	Vazão de fundo	t/h	1,38	2,73

Neste caso, importante observação é a não inclusão, como variáveis de saída, da análise dos componentes Isopreno (IP) e Butino-2 no topo da coluna. Embora seja uma variável importante para operação da coluna, principalmente no que diz respeito à especificação do produto final, não foi possível o ajuste da REDE 9 utilizando estas variáveis. As análises destes componentes apresentam valores praticamente constantes, dificultando o ajuste. Os valores do erro residual para a rede incluindo estas variáveis foram bastante altos e muitas vezes apresentavam um comportamento oscilatório, independente do número de neurônios empregado. Além disso, para efeitos de simulação e otimização, o conjunto de dados para estas variáveis encontra-se fora da faixa desejável de operação e não é possível extrapolação utilizando modelos de redes neurais. Uma vez excluído este conjunto de informações, o erro residual reduziu muito, sendo, então possível o ajuste da rede.

Uma vez definidas as variáveis de entrada e saída e analisados os dados, seguindo a mesma metodologia já descrita anteriormente, estes foram divididos em dois grupos: conjunto de aprendizagem e conjunto de testes e foram, então, realizadas diversas corridas, variando o número de neurônios na camada oculta. Os resultados da variação do erro residual do conjunto de teste em função do número de iterações e neurônios na camada oculta são apresentados no ANEXO A.

O número ótimo de neurônios na camada oculta (NH) correspondente ao menor erro obtido para o conjunto de testes pode ser igual a 15 ou 16. Optou-se por uma rede com 15 neurônios na camada oculta por não se observar diferença significativa nos valores do erro residual para estas configurações. O número de iterações empregado é igual a 50000, uma vez que acima deste valor o ganho na diminuição do erro não é significativo o suficiente para justificar a utilização de um maior número de iterações.

A topologia ou configurações da rede neural selecionada é apresentada na Tabela 4.21 para a REDE 9.

Tabela 4.21. - Configurações da REDE 9

REDE 9	
Número de Pontos do Conjunto de Aprendizagem	163
Número de Pontos do Conjunto de teste	70
Número de Variáveis de Entrada	6
Número de Variáveis de Saída	3
Número de Neurônios na Camada Oculta	15
Número de Neurônios na Camada de Saída	3
Número de Iterações	50000
Fator de Amortecimento	1,00
Parâmetro de Momento	1,00

Similarmente aos casos anteriores, os resultados do subsistema 9 (REDE 9) são apresentados a seguir, de modo a verificar o ajuste do modelo da rede neural. As figuras que apresentam os resultados obtidos no ajuste para o conjunto de dados de teste, comparando os valores das variáveis de saída medidas e calculadas, bem como os testes de avaliação da consistência e desempenho do modelo da rede gerada podem ser vistas no ANEXO B.

De maneira geral, os ajustes podem ser considerados satisfatórios como pode ser visto pelos fatores de correlação.

Os resultados seguintes mostram os testes para verificação da consistência do modelo da rede neural gerada. A partir da distribuição do erro ou função distância, a qual fornece a medida da diferença entre o valor experimental e o calculado pela rede, observa-se que, de modo geral, o ajuste das variáveis pela rede proposta é satisfatório.

A outra forma de verificação da consistência e ajuste dos dados é a representação da distribuição normal dos erros entre os valores experimentais e os gerados pela rede neural, a qual demonstra um bom ajuste do modelo. A distribuição dos erros para todas as variáveis apresenta um comportamento aproximadamente normal e média próxima a zero.

De modo geral, os erros avaliados pelas técnicas mostradas evidenciam um bom ajuste das variáveis modeladas através da REDE 9 proposta, respeitando as considerações observadas para este caso.

4.5.2.10. Subsistema 10 - REDE 10

O subsistema 10 corresponde à Segunda. Coluna de Remoção de Pesados (COL-09), cujo esquema de fluxo para modelagem é apresentado na Fig. 4.31, enquanto que o esquema da rede neural representativa do sistema, REDE 10, com as respectivas variáveis de entrada e saída, é mostrado na Fig. 4.32. As faixas de aplicabilidade destas variáveis podem ser vistas na Tabela 4.22.

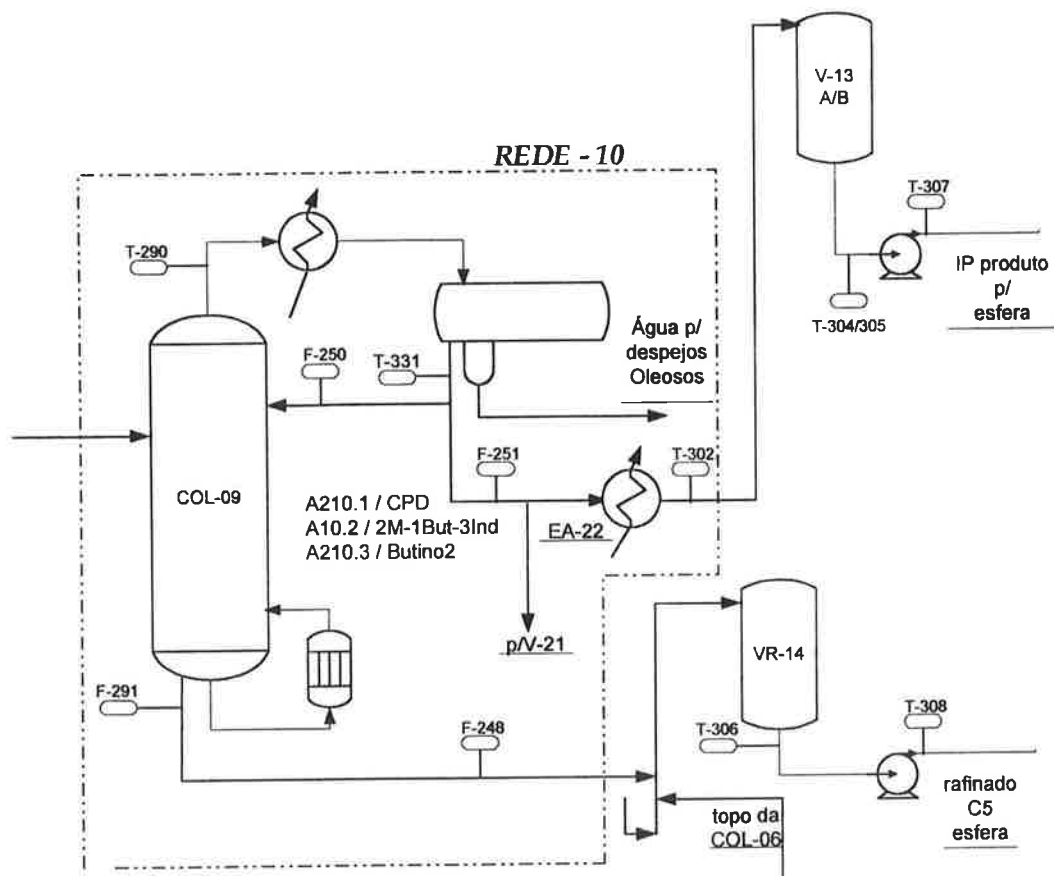


Fig. 4.31. – Fluxograma da COL-09 modelada pela REDE 10

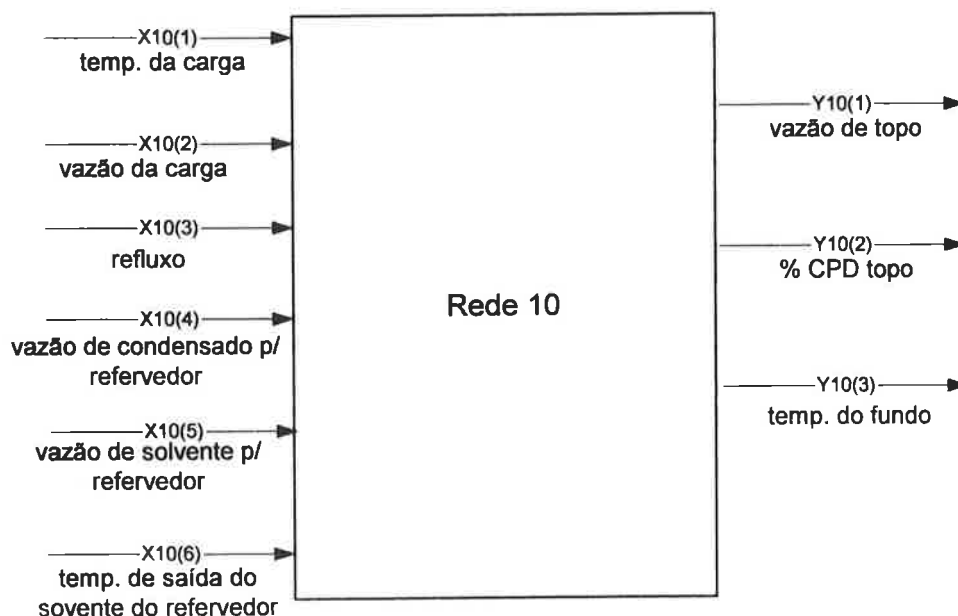


Fig. 4.32 - Esquema da REDE 10 – variáveis de entrada/saída

Tabela 4.22. - Variáveis de Processo para a REDE 10 e Suas Faixas de Aplicação

Notação	Variáveis	Unidades	Valores mínimos	Valores máximos
Entradas				
x10(1)	Temperatura da Carga	(°C)	60,11	63,77
x10(2)	Vazão de Carga	t/h	2,01	2,75
x10(3)	Vazão de Refluxo	t/h	11,39	15,00
x10(4)	Vazão de Condensado p/ Refervedor	t/h	6,95	22,70
x10(5)	Vazão de Solvente p/ Refervedor	t/h	7,85	23,24
x10(6)	Temperatura de Saída do Solvente do Refervedor	(°C)	55,21	66,94
Saídas				
y10(1)	Vazão de topo	t/h	1,86	2,60
y10(2)	Análise de CPD no topo	ppm	0,07	0,52
y10(3)	Temperatura de Fundo	(°C)	49,20	53,59

Neste caso, importante observação é a não inclusão, como variável de saída, da análise do componente Butino-2 no topo da coluna. Embora seja uma variável importante para operação da coluna, principalmente no que diz respeito à especificação do produto final, não foi possível o ajuste da REDE 10 utilizando esta variável. A análise do Butino-2 apresenta um conjunto de dados com alta variabilidade, mas sem qualquer tendência. Os valores do erro residual para a rede incluindo esta variável foram bastante altos e muitas vezes apresentavam um comportamento oscilatório, independente do número de neurônios empregado. Uma vez excluído este conjunto de informações, o erro residual reduziu muito, sendo, então possível o ajuste da rede.

Uma vez definidas as variáveis de entrada e saída e analisados os dados, seguindo a mesma metodologia já descrita anteriormente, estes foram divididos em dois grupos: conjunto de aprendizagem e conjunto de testes e foram feitas diversas corridas variando o número de neurônios na camada oculta. Os resultados da variação do erro residual do conjunto de teste em função do número de iterações e neurônios na camada oculta são apresentados no ANEXO A.

O número ótimo de neurônios na camada oculta (NH) correspondente ao menor erro obtido para o conjunto de teste é igual a 16. O número de iterações empregado corresponde ao erro mínimo e é igual a 26000.

A topologia ou configurações da rede neural selecionada é apresentada na Tabela 4.23 para a REDE 10.

Analogamente aos casos anteriores, os resultados do subsistema 10 (REDE 10) são apresentados a seguir, de modo a verificar o ajuste do modelo da rede neural. As figuras que apresentam os resultados obtidos no ajuste para o conjunto de dados de teste, comparando os valores das variáveis de saída medidas e calculadas, bem como os testes de avaliação da consistência e desempenho do modelo da rede gerada podem ser vistas no ANEXO B.

Tabela 4.23. - Configurações da REDE 10

REDE 10	
Número de Pontos do Conjunto de Aprendizagem	136
Número de Pontos do Conjunto de teste	58
Número de Variáveis de Entrada	6
Número de Variáveis de Saída	3
Número de Neurônios na Camada Oculta	16
Número de Neurônios na Camada de Saída	3
Número de Iterações	26000
Fator de Amortecimento	1,00
Parâmetro de Momento	1,00

De maneira geral, os ajustes podem ser considerados satisfatórios apesar da pequena dispersão apresentada para a variável teor de CPD no topo da COL-09 (Fig. B.28 – ANEXO B), como pode ser visto pelo fator de correlação.

Os resultados seguintes mostram os testes para verificação da consistência do modelo da rede neural gerada. O primeiro teste foi o de distribuição da função distância, a qual fornece a medida da diferença entre o valor experimental e o calculado pela rede. Observa-se que, de modo geral, o ajuste das variáveis pela rede proposta é satisfatório.

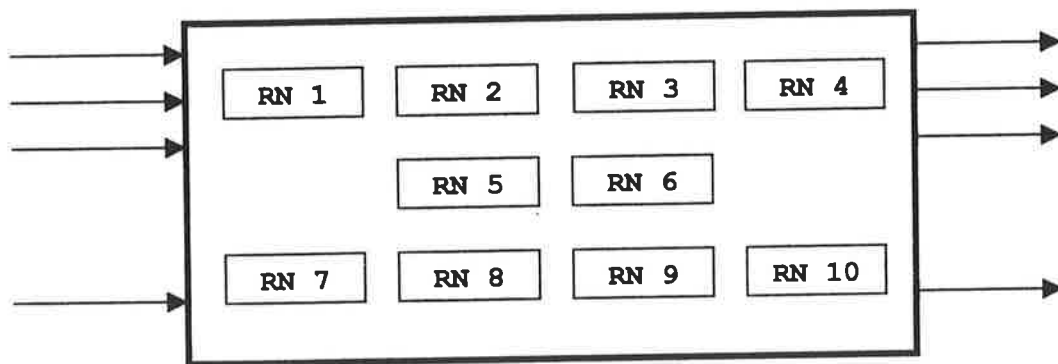
A outra forma de verificação da consistência e ajuste dos dados é a representação da distribuição normal dos erros entre os valores experimentais e os gerados pela rede neural. Os histogramas obtidos mostram um bom ajuste, já que a distribuição dos erros para todas as variáveis apresenta um comportamento aproximadamente normal com média zero.

De modo geral, os erros avaliados pelas técnicas mostradas apresentam evidenciam um bom ajuste das variáveis modeladas através da REDE 10 proposta, respeitando as considerações observadas para este caso.

4.5.3. Modelagem Global do Processo

A metodologia utilizada neste estudo, como já descrito anteriormente, foi baseada na geração de redes individuais representando cada subsistema e, finalmente, interligando estes subsistemas, gerando, desta forma, um modelo global que representa toda a unidade.

Após a modelagem das redes neurais individuais, as quais representam cada subsistema, o principal foco é interconectar seqüencialmente estas redes de modo a gerar dados com o objetivo de construir uma superfície de resposta para todas as variáveis envolvidas, bem como analisar o comportamento das variáveis de processos mais importantes. Assim uma “caixa preta” foi construída, como mostrado no exemplo da Fig. 4.1. no item 4.2.2. sobre modelagem e ajuste do sistema, contendo todas as redes interconectadas por correntes seqüenciais e recícos. Esta figura é mostrada uma vez mais a seguir.



O procedimento de convergência do sistema integrado foi executado interconectando todas as correntes seqüencialmente às correspondentes operações unitárias ou subsistemas, de modo a reproduzir a planta industrial. O algoritmo de substituição direta com fator de amortecimento foi usado para convergência matemática.

O modelo global é complexo devido ao grande número de variáveis e ciclos envolvidos, no entanto tanto o modelo gerado e a estrutura de convergência empregada mostram boa eficiência na representação do sistema.

A Fig. 4.2., no item 4.4.1, mostra, esquematicamente, a Unidade de Produção de Isopreno e a seguir serão apresentados, mais detalhadamente, diagramas de fluxos para os diversos subsistemas, incluindo a instrumentação, as variáveis em comum e os ciclos, para melhor entendimento das conexões entre eles.

A Fig. 4.33 mostra os subsistemas constituintes da Seção de Preparação de Carga ou Sistema de Concentração de Isopreno.

As Figs. 4.34 a 4.36 mostram os subsistemas constituintes dos Sistemas de Destilação Extrativa e Lavagem e de Circulação e Regeneração de Solvente.

A Fig. 4.37 mostra os subsistemas constituintes da Seção de Fracionamento.

A estrutura para modelagem e simulação global do processo, mostrando as conexões entre os diversos subsistemas pode ser visto na Fig. 4.38. A seqüência de cálculo das redes utilizada para modelagem foi: REDE 1, REDE 2, REDE 3, REDE 6, REDE 4, REDE 5, REDE 7, REDE 8, REDE 9 e REDE 10. A convergência foi realizada apenas para a variável $x_4(5)$. Utilizou-se um programa desenvolvido, no Laboratório de Simulação e Controle de Processos (LSCP) do Departamento de Engenharia Química da Universidade de São Paulo, por Nascimento e Alves (2003).

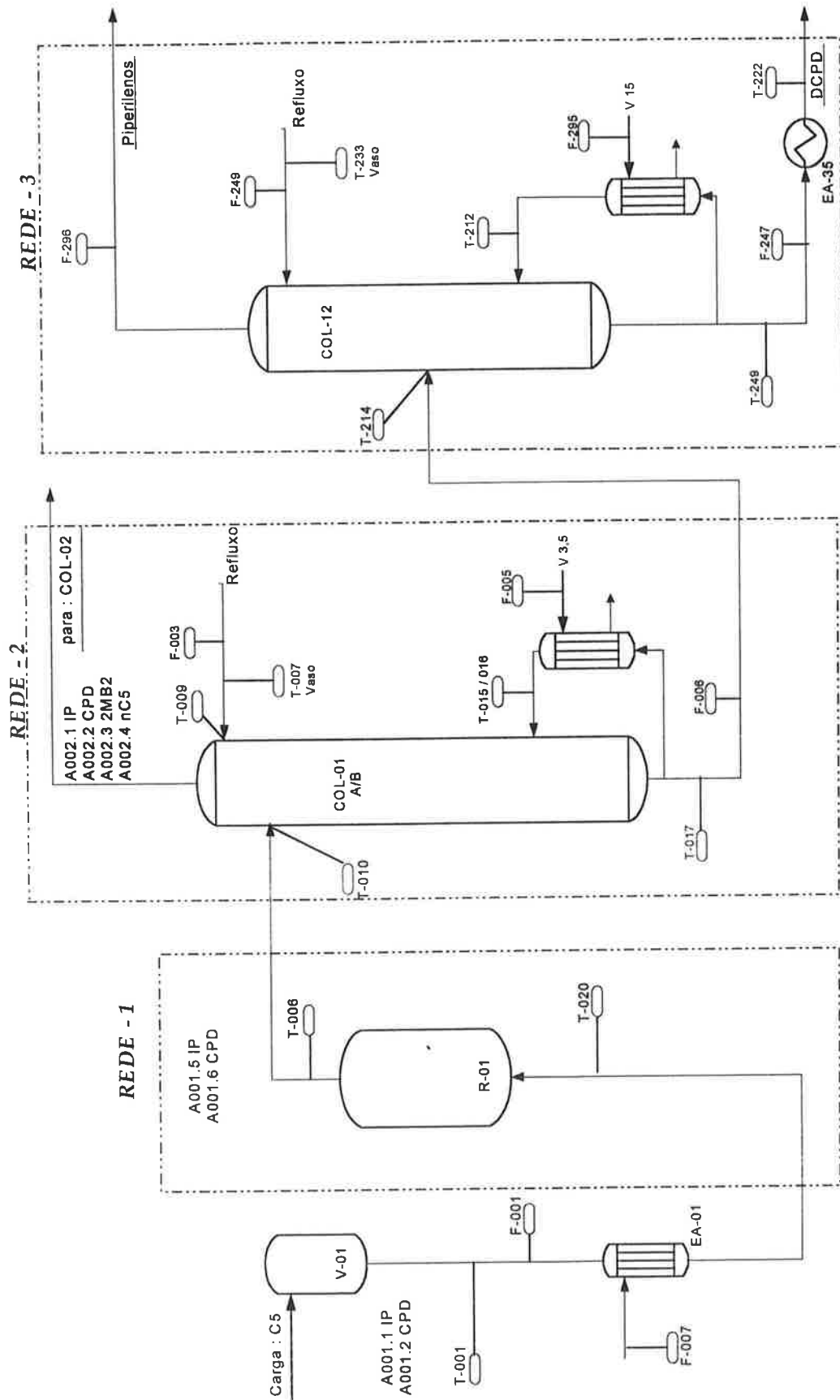


Fig. 4.33. – Representação de Fluxo da Seção de Preparação de Carga

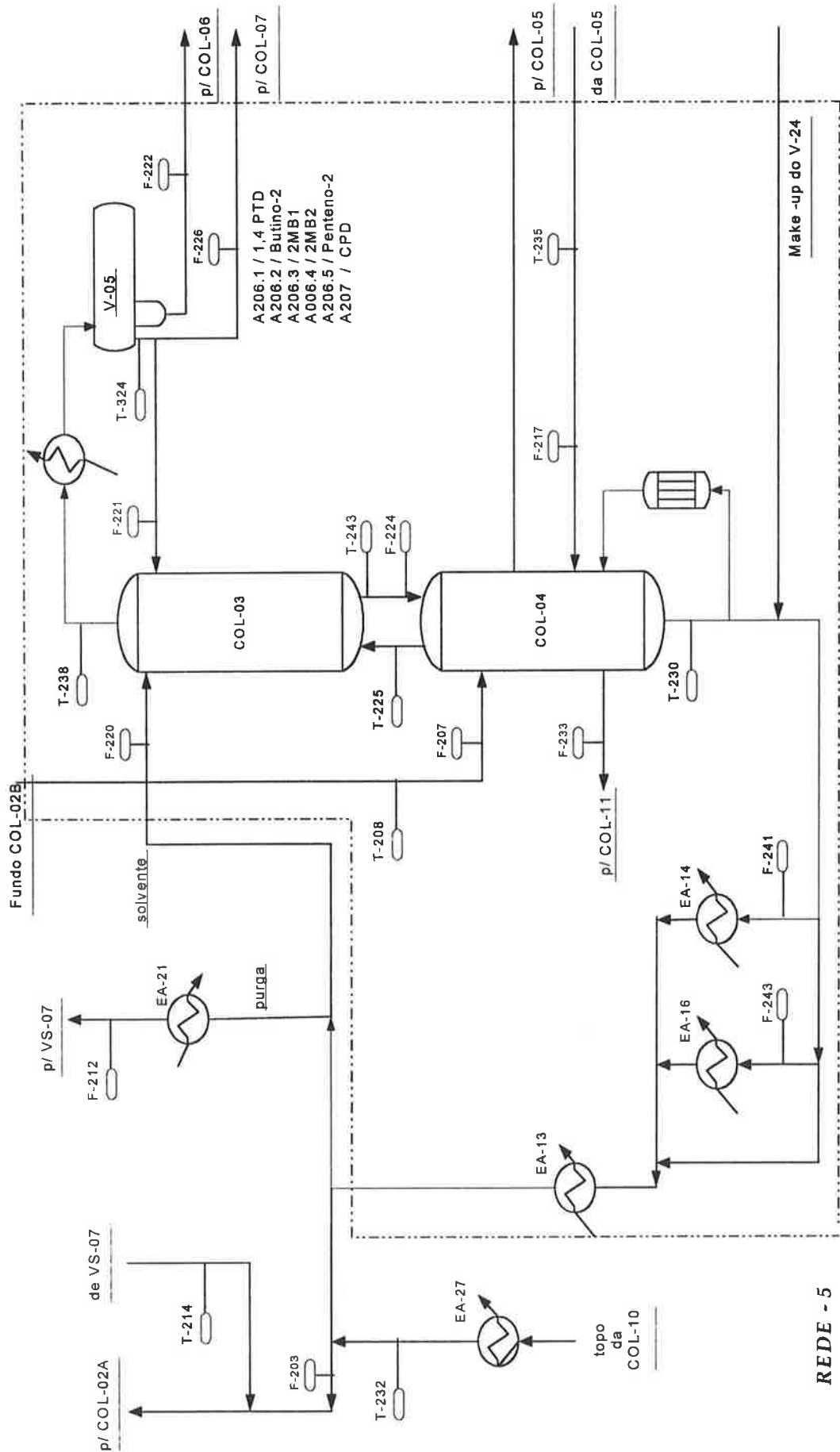


Fig. 4.35 – Sistema de Extração – Segunda Coluna de Destilação Extrativa e Stripper de Solvente

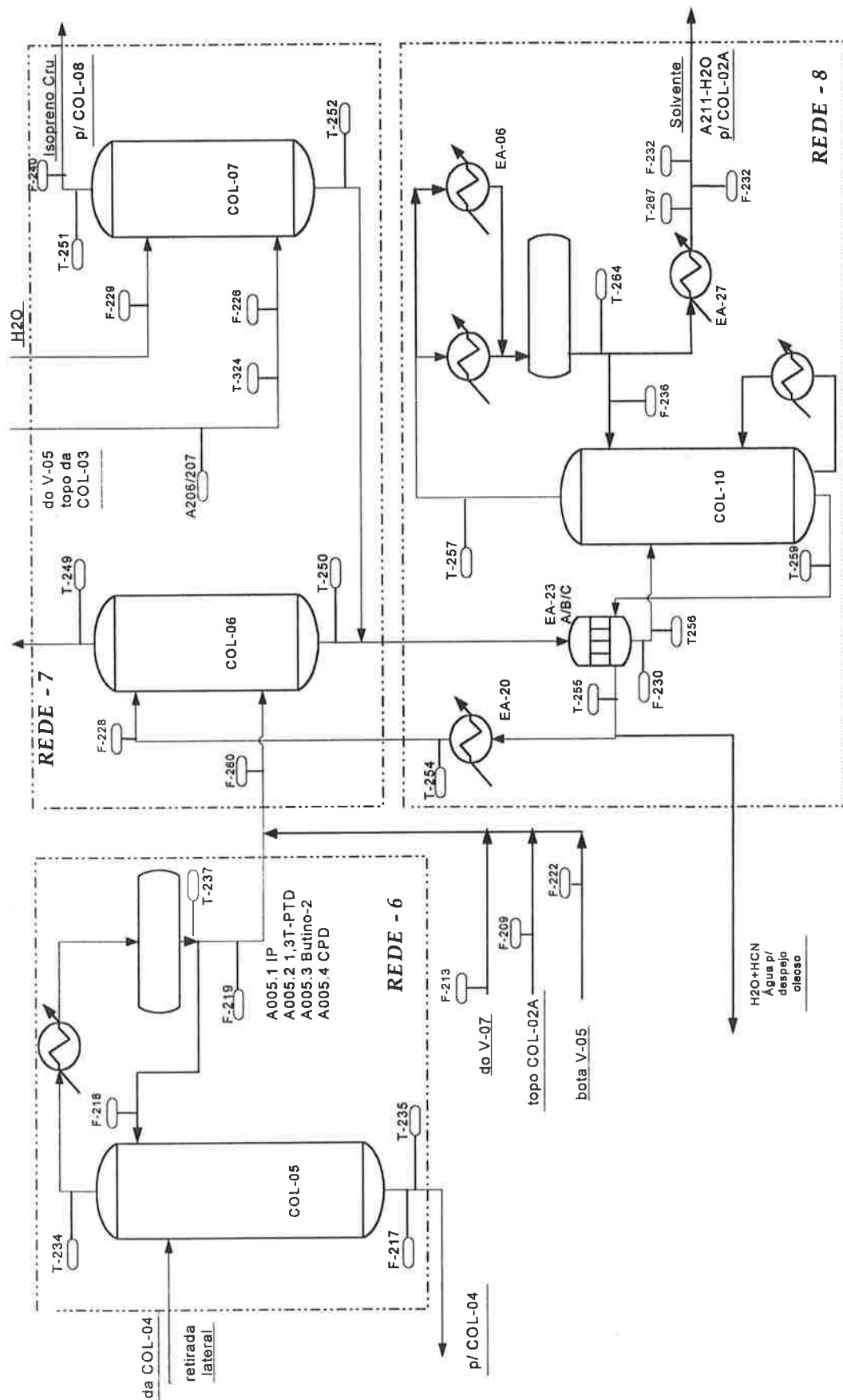


Fig. 4.36. – Sistema de Lavagem e Recuperação de Solvente

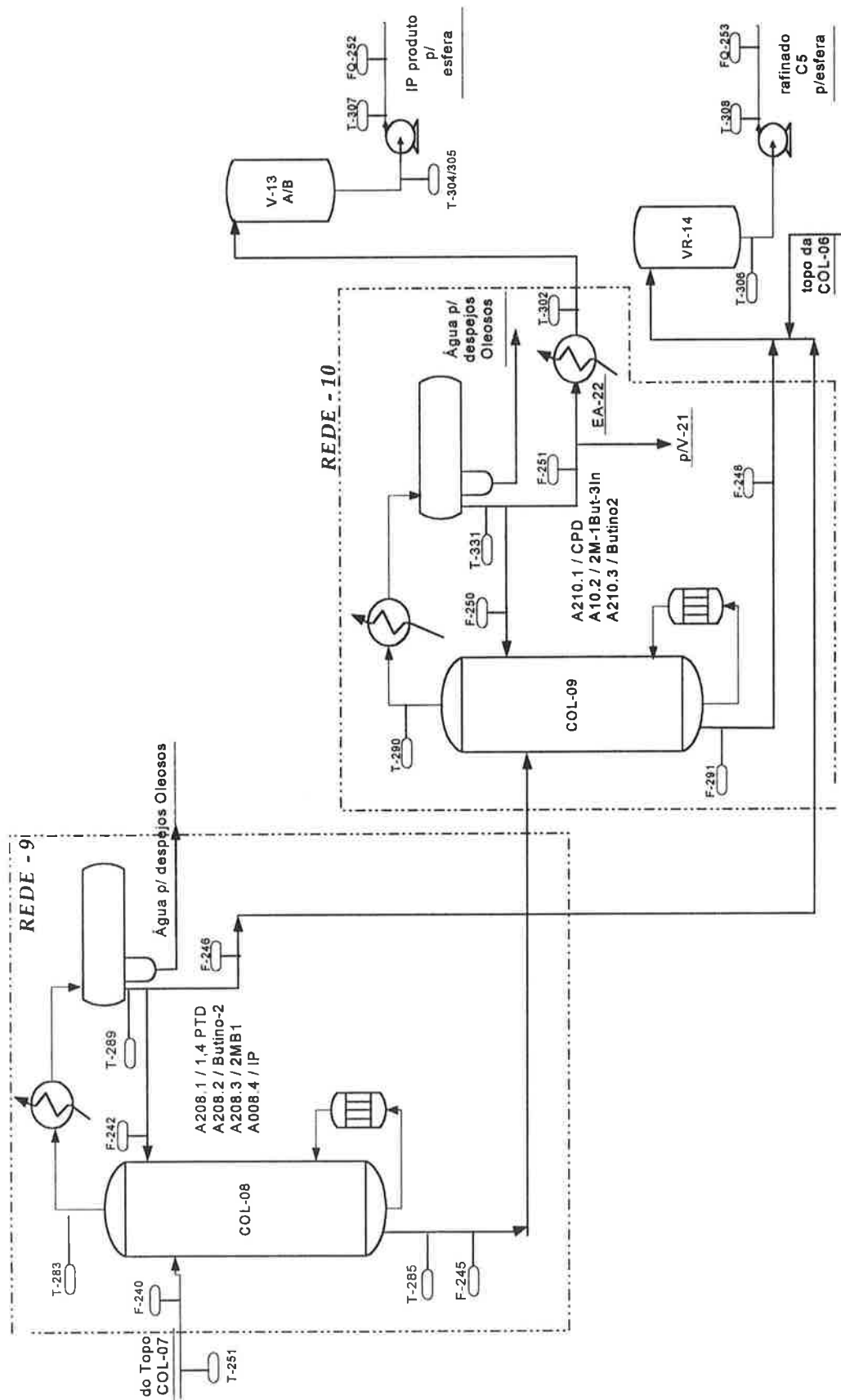


Fig. 4.37 – Seção de Fracionamento

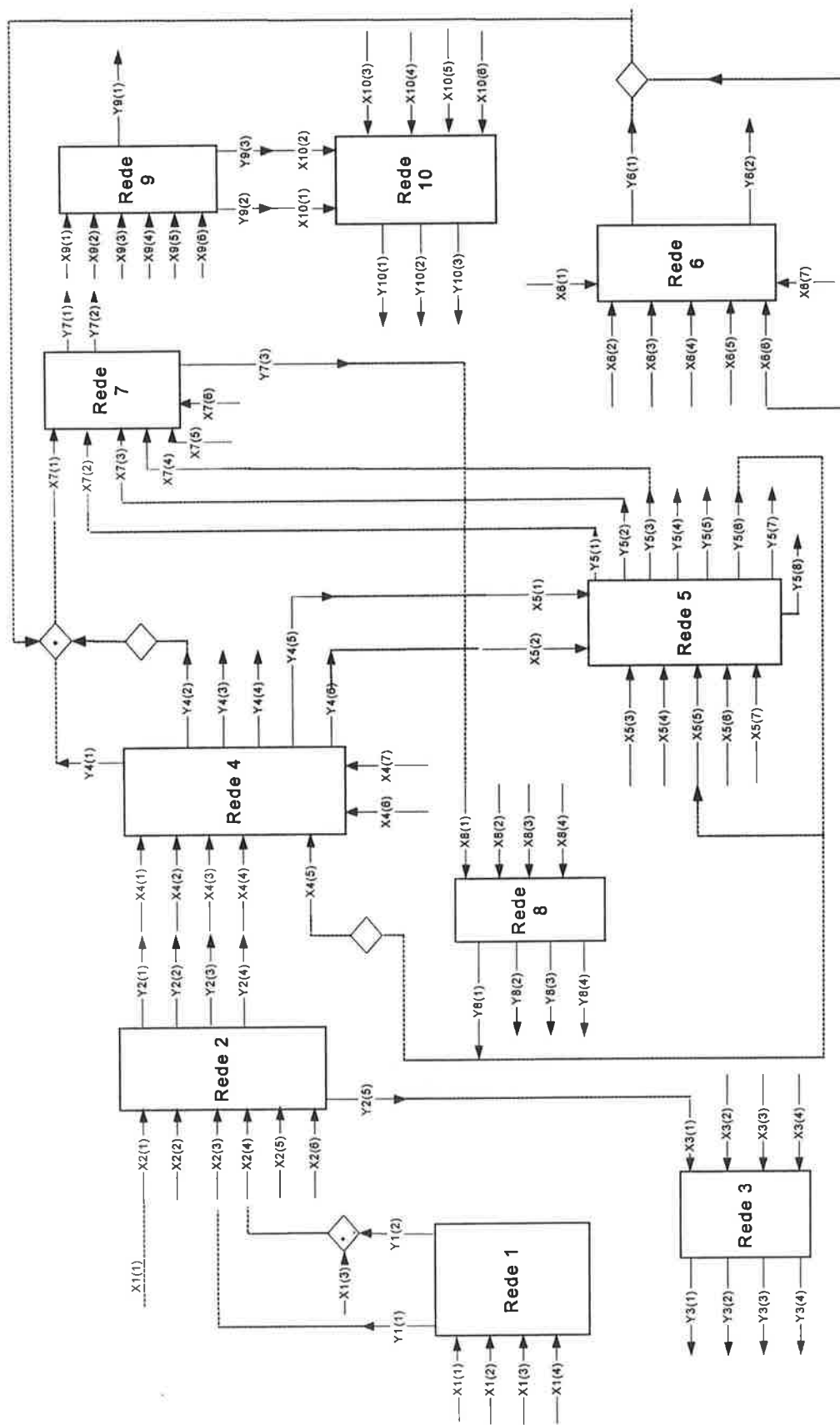


Fig. 4.38. - Esquema de Modelagem usando Redes Neurais – Simulação do Processo Global.

A Tabela 4.24 resume todas as variáveis consideradas nesta simulação, suas igualdades e descrições

Tabela 4.24 – Variáveis Modeladas para o Processo de Produção de Isopreno

Notação	Descrição da Variável	Igualdades
REDE 1		
x1(1)	Vazão de carga	
x1(2)	Análise de IP na carga	
x1(3)	Análise de CPD na carga	
x1(4)	Temperatura da Carga	
y1(1)	Análise de IP no reator	
y1(2)	CPD carga - CPD reator	
REDE 2		
x2(1)	Vazão de carga	$x2(1) = x1(1)$
x2(2)	Temperatura da Carga	
x2(3)	Análise de IP na carga	$x2(3) = y1(1)$
x2(4)	Análise de CPD na carga	$x2(4) = x1(3) - y1(2)$
x2(5)	Vazão de Refluxo	
x2(6)	Vapor para Refervedor	
y2(1)	Vazão de topo	
y2(2)	Análise de IP no topo	
y2(3)	Análise de CPD no topo	
y2(4)	Análise de 2M2B no topo	
y2(5)	Vazão de fundo	
REDE 3		
x3(1)	Vazão de carga	$x3(1) = y2(5)$
x3(2)	Vazão de Refluxo	
x3(3)	Temperatura da carga	
x3(4)	Vapor para Refervedor	
y3(1)	Vazão de topo	
y3(2)	Temperatura de topo	
y3(3)	Vazão de fundo	
y3(4)	Temperatura de fundo	
REDE 4		
x4(1)	Vazão de carga	$x4(1) = y2(1)$
x4(2)	Análise de IP na carga	$x4(2) = y2(2)$
x4(3)	Análise de CPD na carga	$x4(3) = y2(3)$

Tabela 4.24 (cont.) – Variáveis Modeladas para o Processo de Produção de Isopreno

Notação	Descrição da Variável	Igualdades
x4(4)	Análise de 2M2B na carga	$x4(4) = y2(4)$
x4(5)	Vazão de Solvente	$x4(5) = y8(1) + y5(6) - x5(5) + 0,20$
x4(6)	Vazão de Refluxo	
x4(7)	Vapor para Refervedor	
y4(1)	Vazão de topo I	
y4(2)	Vazão de topo II	
y4(3)	Análise de i-C5 no topo	
y4(4)	Análise de n-C5 no topo	
y4(5)	Vazão de fundo	
y4(6)	Temperatura de fundo	
REDE 5		
x5(1)	Vazão de Carga I	$x5(1) = y4(5)$
x5(2)	Temperatura de Carga I	$x5(2) = y4(6)$
x5(3)	Vazão de Carga II	$x5(3) = x6(6)$
x5(4)	Temperatura de Carga II	$x5(4) = x6(7)$
x5(5)	Vazão de Solvente	
x5(6)	Vazão de Refluxo	
x5(7)	Vapor para Refervedor	
y5(1)	Vazão de Topo	
y5(2)	Análise de Butino-2 no topo	
y5(3)	Análise de CPD no topo	
y5(4)	Retirada Lateral	$y5(4) = y6(1)$
y5(5)	Temperatura # 55	
y5(6)	Vazão de Fundo	
y5(7)	Temperatura de Fundo	
y5(8)	Análise de H2O no Solvente (fundo da Coluna 04)	
REDE 6		
x6(1)	Temperatura de Topo	
x6(2)	Temperatura do Vaso de Topo	
x6(3)	Análise de IP no Topo	
x6(4)	Análise de Butino-2 no Topo	
x6(5)	Análise de CPD no Topo	
x6(6)	Vazão de Fundo	
x6(7)	Temperatura de Fundo	
y6(1)	Vazão de Carga	
y6(2)	Vazão de Refluxo	

Tabela 4.24 (cont.) – Variáveis Modeladas para o Processo de Produção de Isopreno

Notação	Descrição da Variável	Igualdades
REDE 7		
x7(1)	Vazão de Carga I	$x7(1) = y4(1) + y4(2) + 0,2 + y6(1) - x6(6)$
x7(2)	Vazão de Carga II	$x7(2) = y5(1)$
x7(3)	Análise de Butino-2 na Carga	$x7(3) = y5(2)$
x7(4)	Análise de CPD na Carga	$x7(4) = y5(3)$
x7(5)	Vazão de Água I	
x7(6)	Vazão de Água II	
y7(1)	Vazão de Topo (IP cru)	
y7(2)	Temperatura de Topo	
y7(3)	Vazão de Fundo	
REDE 8		
x8(1)	Vazão de Carga	$x8(1) = y7(3)$
x8(2)	Temperatura de Carga	
x8(3)	Vazão de Refluxo	
x8(4)	Vapor para Refervedor	
y8(1)	Vazão de topo	
y8(2)	Temperatura de topo	
y8(3)	Análise de H2O no topo	
y8(4)	Temperatura # 10	
REDE 9		
x9(1)	Vazão de carga	$x9(1) = y7(1)$
x9(2)	Temperatura de Carga	$x9(2) = y7(2)$
x9(3)	Vazão de Refluxo	
x9(4)	Vazão de Solvente p/ Refervedor	
x9(5)	Temperatura de Entrada do Solvente no Refervedor	
x9(6)	Temperatura de Saída do Solvente do Refervedor	
y9(1)	Temperatura do vaso	
y9(2)	Temperatura de Fundo	
y9(3)	Vazão de fundo	
REDE 10		
x10(1)	Temperatura da Carga	$x10(1) = y9(2)$
x10(2)	Vazão de Carga	$x10(2) = y9(3)$
x10(3)	Vazão de Refluxo	
x10(4)	Vazão de Condensado p/ Refervedor	
x10(5)	Vazão de Solvente p/ Refervedor	
x10(6)	Temperatura de Saída do Solvente do Refervedor	

Tabela 4.24 (cont.) – Variáveis Modeladas para o Processo de Produção de Isopreno

Notação	Descrição da Variável	Igualdades
y10(1)	Vazão de topo	
y10(2)	Análise de CPD no topo	
y10(3)	Temperatura de Fundo	

4.5.4. Simulação do Processo

Uma vez gerado o modelo global processo, torna-se necessário, validar o modelo de modo a poder aplicá-lo em estudos de sensibilidade, determinação das restrições do processo, otimização, entre outros. A validação do modelo passa por simulações do sistema, comparando os resultados calculados com os dados da planta industrial ou com os resultados gerados por modelos rigorosos, utilizando simuladores comerciais. Neste trabalho, foram empregados dados do processo industrial. Arquivos com variáveis de entrada dadas aleatoriamente, restritas a faixa de domínio de cada uma, foram fornecidos como entrada para o programa de simulação, o qual convergiu para valores dentro dos domínios das variáveis de saída, demonstrando o poder de interpolação das redes neurais. Embora o modelo global gerado seja complexo devido ao grande número de variáveis e ciclos existentes, este se mostrou bastante eficaz para o presente estudo.

As Tabelas apresentando os conjuntos de dados reais da planta industrial para a avaliação da convergência do sistema para alguns dos casos estudados e os resultados da simulação do processo global, para estes casos, são mostrados no ANEXO C.

4.5.5. Análise de Sensibilidade

A análise de sensibilidade das variáveis de um processo constitui uma das etapas fundamentais para o conhecimento operacional global do processo e, conseqüentemente, é um dos passos importantes para um melhor entendimento da interação entre as variáveis manipuladas e controladas, isto é, a relação entre as variáveis dependentes e independentes do processo. Assim, após a validação do modelo gerado com os dados da planta, foram realizados estudos de sensibilidade para algumas das variáveis consideradas importantes para o processo de produção de Isopreno. Serão apresentados nesta seção, exemplos de resultados deste estudo, de modo a demonstrar a potencialidade do uso de simuladores com modelos de redes neurais neste tipo de análise.

Optou-se por mostrar os resultados da Seção de Preparação de Carga devido a sua importância na estabilidade e especificação de variáveis chaves para as demais seções da unidade de produção de isopreno.

Para um melhor entendimento, será feita uma breve descrição do processo, de modo a analisar as variáveis de processo manipuladas (ou perturbadas) e seus efeitos nas variáveis de interesse e nas especificações desejadas.

Como descrito anteriormente, a seção de preparação de carga tem por objetivo a concentração de isopreno (IP) na corrente C_5 que alimenta a seção de destilação extrativa da unidade, além de garantir a especificação do ciclopentadieno (CPD) nesta corrente. Como os sub-produtos comerciais desta unidade (Piperilenos e Diciclopentadieno, DCPD) são obtidos nesta seção, é desejável, que o CPD possa ser especificado no reator R-01, liberando a coluna COL-01 para especificar o n-pentano ($n-C_5$) na corrente rica em piperilenos.

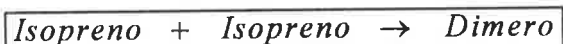
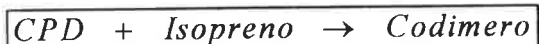
Devido à impossibilidade de modelar todas as variáveis necessárias para uma análise mais aprofundada desta seção, o presente estudo enfoca basicamente, e com mais rigor, a modelagem do reator e em seguida aborda, sucintamente,

variáveis relevantes para a operação das colunas COL-01 e COL12, de modo a especificar os subprodutos da unidade e a carga da seção de extração.

4.5.5.1. Análise de Sensibilidade do Reator

A necessidade de um reator neste processo deve-se ao fato da volatilidade relativa do CPD ser muito próxima à do IP, o que torna praticamente inviável a separação destes compostos por destilação. O reator R-01 favorece a reação térmica do CPD em um composto mais pesado DCPD (diciclopentadieno), facilitando desta maneira a separação do isopreno. As reações que ocorrem no R-01 são homogêneas, em fase líquida, não catalíticas e exotérmicas. O reator é isolado termicamente e possui fluxo ascendente. A conversão do CPD à DCPD é espontânea, mesmo à temperatura ambiente.

As principais reações que ocorrem no reator são:



Além destas reações, podem ocorrer outras envolvendo os demais dienos, como por exemplo os piperilenos, com o próprio CPD ou isopreno. Poucas informações existem a respeito da conversão destas reações. Observando dados operacionais das composições do CPD, isopreno e piperilenos na entrada e saída do reator conclui-se que as reações envolvendo os piperilenos podem ser desprezadas na modelagem.

Além das reações de dimerizações comentadas anteriormente, também podem ocorrer no DC-5201 reações de polimerização, com formação de trímeros, tetrâmeros, etc. A polimerização do CPD é favorecida em condições de temperatura superiores a 100°C. A reação de dimerização do CPD predomina em relação às demais (Pellicci e Carneiro, 1996).

À priori, sabe-se que é importante o controle da temperatura da carga de forma preditiva. A temperatura da carga do reator é função da vazão e das concentrações de CPD e IP na carga, sendo o teor de CPD é a variável mais importante para o controle operacional da temperatura de entrada do mesmo de modo a garantir a especificação de 3% de CPD na saída do reator. Este valor é suficiente para especificar este composto na carga da extração, sem fazer uso da COL-01 para esta finalidade, podendo, portanto, ficar liberada para especificar o n-C5 na corrente rica em piperileno. Assim, para este estudo, foram gerados dados resultantes da simulação de uma carga típica real (caso base), e, em seguida, variando-se a composição de CPD e a temperatura da carga. Foram analisadas também a influência do teor de IP e a vazão da carga. A Tabela 4.25 mostra os resultados para o caso considerado como base na modelagem do reator (REDE 1). A partir deste caso típico, o teor de CPD, variável x1(3), e a temperatura da carga, variável x1(4), foram variados dentro da faixa de domínio destas variáveis, mantendo-se constante a vazão da carga, variável x1(1), e a composição de IP na carga, variável x1(2). Os resultados podem ser vistos nas Figuras 4.39 e 4.40, para o CPD na saída e perda de IP respectivamente.

Tabela 4.25 – Caso Base para Análise de Sensibilidade

Entrada		Saída	
x1(1)	14,02	y1(1)	14,30
x1(2)	14,81	y1(2)	1,33
x1(3)	4,35		
x1(4)	71,17		

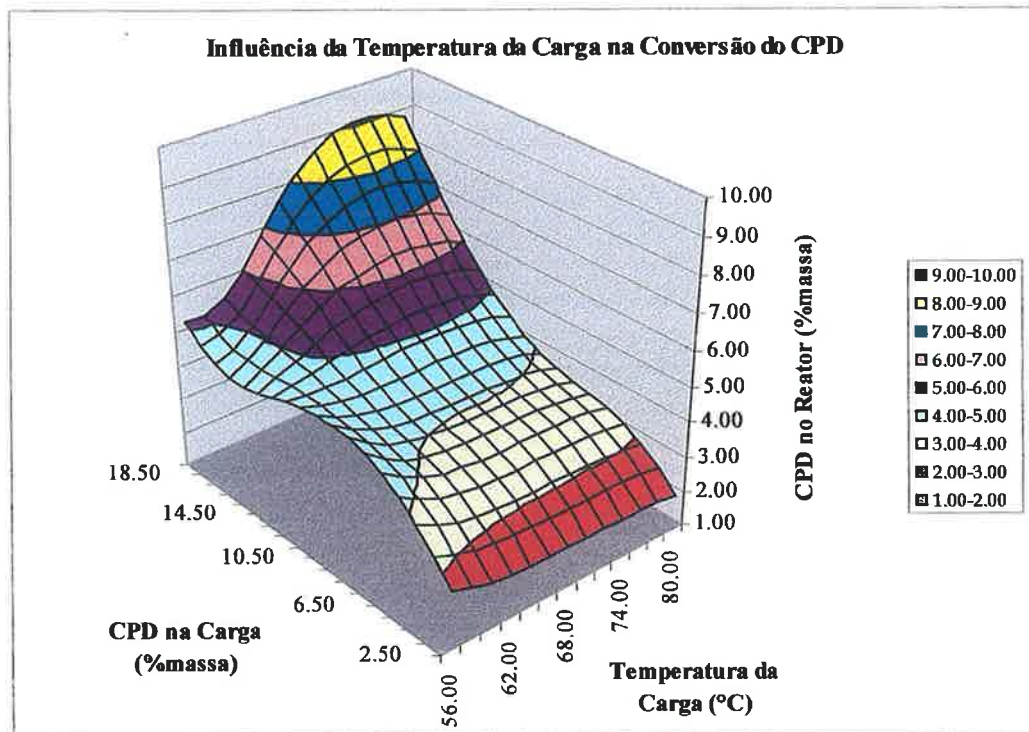


Fig. 4.39. – Influência da Temperatura da Carga do Reator na Conversão de CPD
 Caso 1 - Vazão de Carga: 14,02 e %IP na Carga: 14,81

Observa-se que para um mesmo teor de CPD na carga até aproximadamente 11,50, quanto maior a temperatura da carga, maior a conversão de CPD no reator, isto é, menor o teor de CPD na saída do reator. Para valores de CPD na carga na faixa aproximadamente entre 11,50 e 12,50, a conversão de CPD no reator aumenta até uma dada temperatura, a partir da qual existe uma faixa de temperatura, na qual a conversão diminui e depois volta a aumentar, isto é, o teor de CPD na saída do reator diminui, apresenta um máximo e volta a diminuir. Para valores acima de 13,50, quanto maior a temperatura, menor a conversão de CPD, ou seja, maior o teor de CPD na saída do reator, apresentando um comportamento inverso ao apresentado para baixos teores de CPD na carga.

Para um mesmo valor de temperatura, o teor de CPD na saída do reator apresenta um máximo e depois volta a diminuir com o aumento do teor de CPD na carga. Existe um valor de temperatura acima de 60°C, para o qual o teor de CPD

no reator sempre aumenta com o aumento do teor de CPD na carga. Para teor de CPD igual a 2,50%, os valores válidos de temperatura correspondem àqueles maiores que 60°C, de acordo as restrições impostas pelo domínio das variáveis.

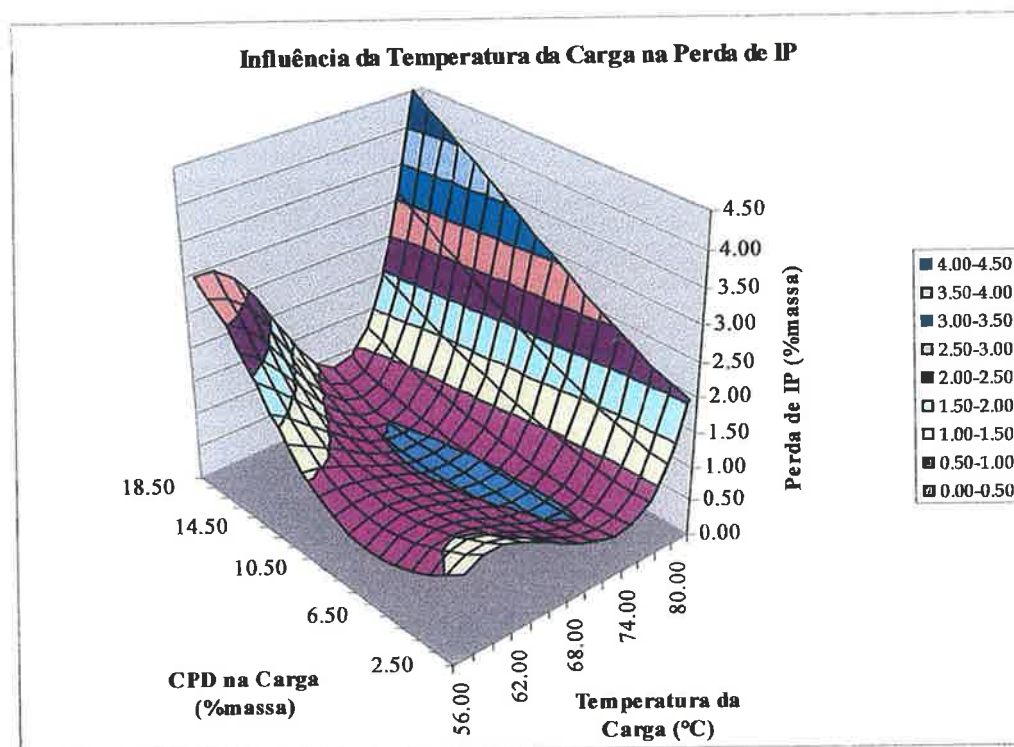


Fig. 4.40. – Influência da Temperatura da Carga do Reator na Perda de IP no Reator
Caso 1 - Vazão de Carga: 14,02 e %IP na Carga: 14,81

A Fig. 4.40 mostra que os valores mínimos de perda de IP podem ser obtidos em temperaturas aproximadamente iguais a 70 – 72° C, sendo que os maiores valores de perda de IP correspondem aos maiores teores de CPD na carga para uma mesma temperatura, atingindo um máximo para maiores temperaturas e maiores teores de CPD na carga. O menor valor de perda de IP ocorre para valores de CPD na carga entre 10,5 e 11,5 e temperatura na carga de aproximadamente 70°C.

A etapa seguinte foi variar a composição de IP na carga do reator entre os valores limites do domínio desta variável. Foram escolhidos, aleatoriamente os seguintes valores 12,00%, 17,50% e 20,00%. Para cada um destes valores, o

mesmo procedimento acima foi adotado e os resultados podem ser vistos a seguir. As Figuras 4.41 e 4.42 apresentam os resultados para 12,00% de IP na carga do reator. A Fig. 4.41 apresenta os resultados para o teor de CPD na saída do reator.

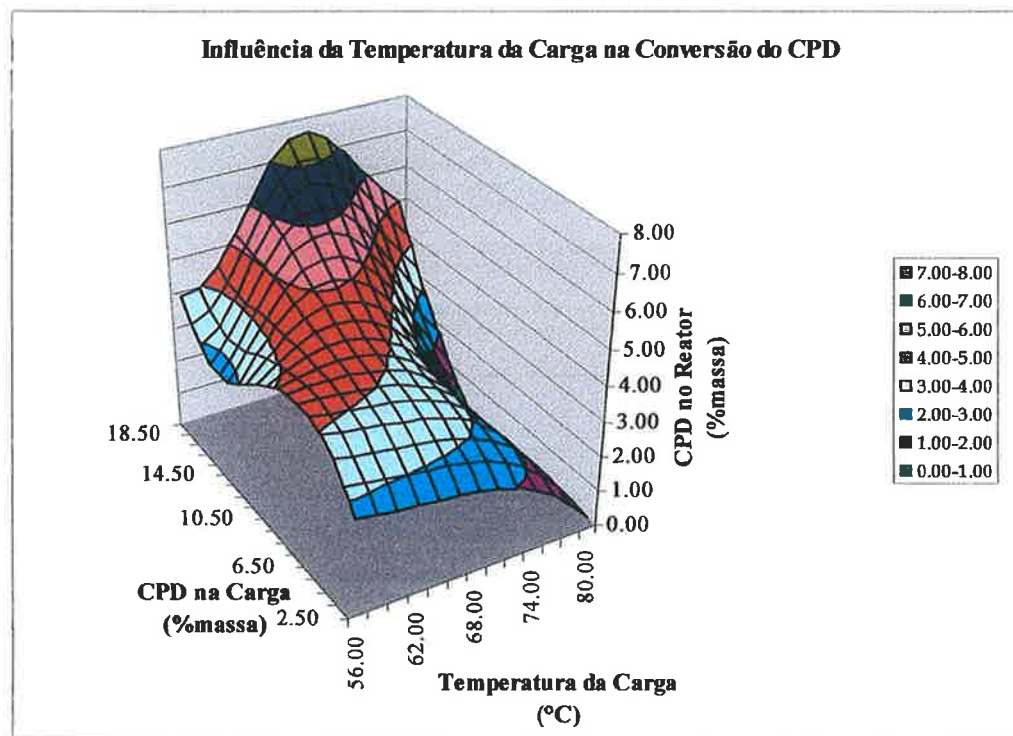


Fig. 4.41. – Influência da Temperatura da Carga do Reator na Conversão de CPD
Caso 2 - Vazão de Carga:14,02 e %IP na Carga: 12,00

Verifica-se que para um mesmo teor de CPD na carga, quanto maior a temperatura da carga, maior a conversão de CPD, isto é, menor o teor de CPD na saída do reator. Para teores de CPD acima de 8,50, existe uma faixa de temperatura na qual a conversão de CPD é mínima, isto é, o teor de CPD na saída do reator é máximo, voltando a diminuir para valores maiores de temperatura. Para um mesmo valor de temperatura, o teor de CPD na saída atinge um máximo e volta a diminuir até a temperatura de aproximadamente 62°C, quando quanto maior o teor de CPD na carga, o teor de CPD na saída do reator sempre aumenta.

A Fig. 4.42 apresenta os resultados para a perda de IP no reator. Observa-se que para um mesmo teor de CPD na carga, a perda de IP apresenta um mínimo em aproximadamente 60°C e aumenta a partir de aproximadamente 62°C. Para um mesmo valor de temperatura, para temperaturas menores que 68°C, a perda de IP apresenta um mínimo para teores intermediários de CPD na carga. Para valores de temperaturas acima deste valor, a perda de IP aumenta à medida que aumenta o teor de CPD na carga e quanto maior a temperatura a perda tende a um valor constante.

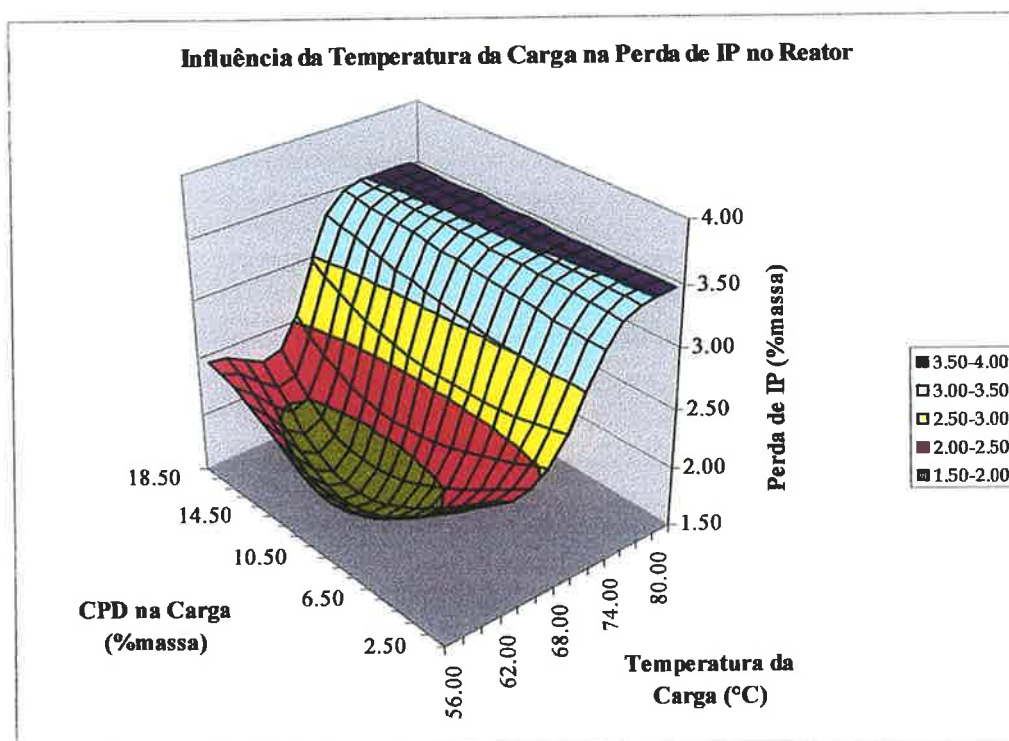


Fig. 4.42. – Influência da Temperatura da Carga do Reator na Perda de IP no Reator
 Caso 2 - Vazão de Carga: 14,02 e %IP na Carga: 12,00

As Figuras 4.43 e 4.44 apresentam os resultados para 17,50% de IP na carga do reator. A Fig. 4.43 apresenta os resultados para o teor de CPD na saída do reator.

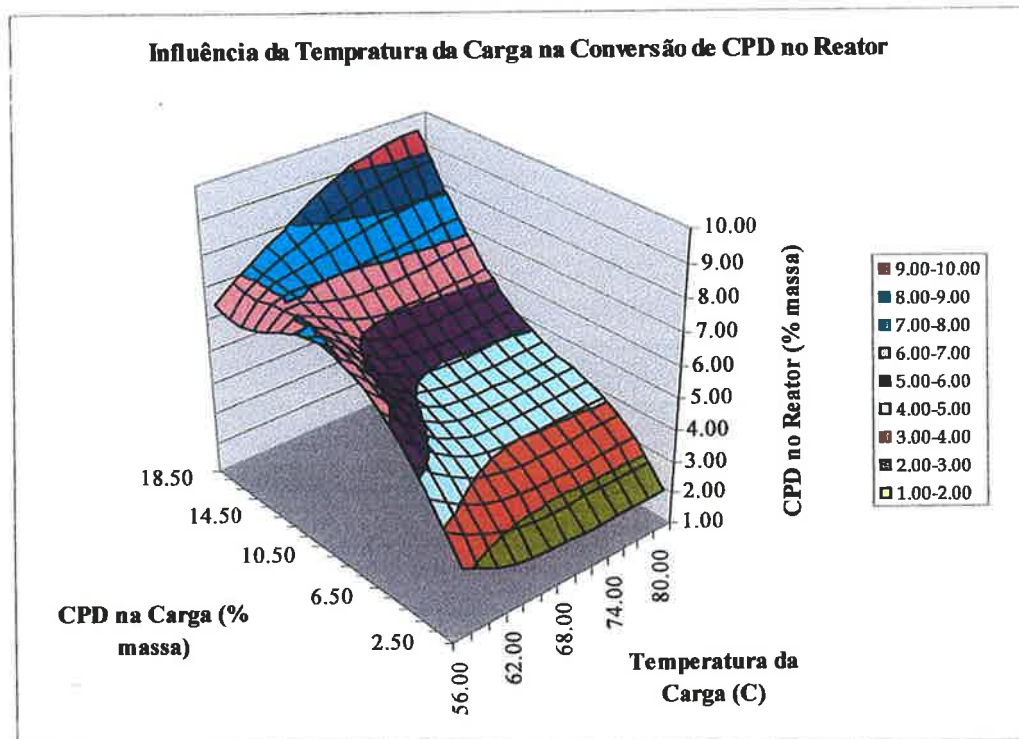


Fig. 4.43. – Influência da Temperatura da Carga do Reator na Conversão de CPD
Caso 3 - Vazão de Carga: 14,02 e %IP na Carga: 17,50

Para um mesmo valor de CPD na carga, o teor de CPD na saída do reator diminui à medida que a temperatura da carga também aumenta, isto é, a conversão aumenta, até atingir um valor mínimo, correspondendo à conversão máxima. A partir deste valor mínimo, o teor de CPD na carga volta a aumentar com o aumento da temperatura. Para valores de CPD na carga de até aproximadamente 13,50, a temperatura correspondente à máxima conversão é cerca de 72°C. Para maiores valores de CPD na carga (>13,50%), a temperatura correspondente à máxima conversão diminui à medida que o teor de CPD na carga aumenta.

Para um mesmo valor de temperatura na carga, para baixas temperaturas, o teor de CPD na saída do reator aumenta até um valor máximo para valores intermediários de CPD na carga (11,50 – 14,50%), quando, em seguida, diminui. Para valores de temperatura maiores que 60°C, o teor de CPD na saída do reator sempre aumenta com o aumento do teor de CPD na carga.

Os maiores valores de CPD na saída correspondem aos maiores valores de CPD e temperatura na carga e os menores valores de CPD na saída (desejável) correspondem a baixos teores de CPD (~ 4,50%) e temperaturas entre 62 e 72°C.

A Fig. 4.44 apresenta os resultados para a perda de IP no reator.

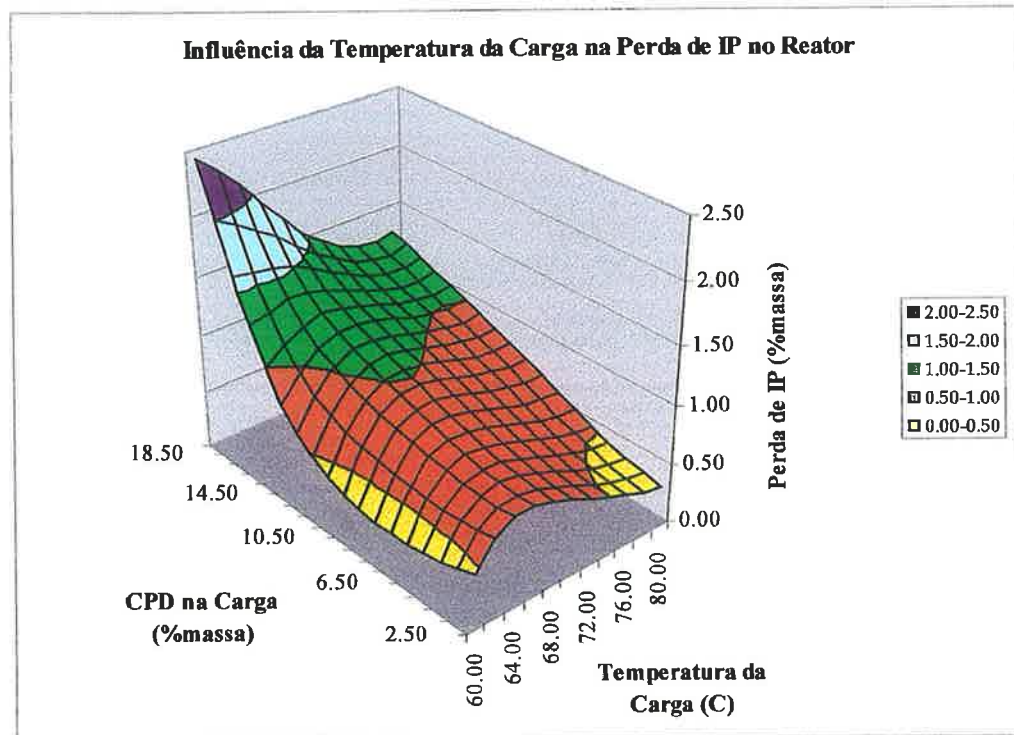


Fig. 4.44. – Influência da Temperatura da Carga do Reator na Perda de IP no Reator
Caso 3 - Vazão de Carga: 14,02 e %IP na Carga: 17,50

Para um mesmo valor de CPD na carga e temperatura acima de 60°C, a perda de IP aumenta com o aumento da temperatura até aproximadamente 68°C. A partir deste valor a perda de IP diminui.

Para um mesmo valor de temperatura na carga, a baixas temperaturas, a perda de IP diminuir com o aumento do teor de CPD na carga até um mínimo (quanto maior a temperatura, menor o teor de CPD na carga necessário para atingir este mínimo). A partir deste mínimo, a perda de IP aumenta com o aumento do teor de CPD na carga. Para valores de temperatura acima de 66°C, para um mesmo

valor de temperatura, quanto maior o teor de CPD na carga, maior a perda de IP no reator.

Menores valores de perda de IP correspondem a baixas temperaturas e baixos valores de CPD na carga ou a altas temperaturas e baixos valores de CPD na carga (região amarela da Fig. 4.44). Neste caso, valores ótimos de temperatura, para obtenção de teores baixos de CPD na saída do reator e menores perdas de IP, correspondem a baixos valores de temperatura.

As Figuras 4.45 e 4.46 apresentam os resultados para 20,00% de IP na carga do reator. A Fig. 4.45 apresentam os resultados para o teor de CPD na saída do reator.

A partir da Fig. 4.45, verifica-se que para um mesmo valor de CPD na carga, o teor de CPD na saída do reator diminui com o aumento da temperatura até atingir um valor mínimo (máxima conversão) em torno de 72°C, temperatura a partir da qual, observa-se um aumento no teor de CPD na saída do reator. Para valores de CPD na carga acima de 11,50%, o teor de CPD aumenta até cerca de 60°C e a partir desta temperatura, o comportamento verificado é idêntico ao já mencionado.

Para um mesmo valor de temperatura, para baixas temperaturas, até aproximadamente 60°C, à medida que aumenta o teor de CPD na carga, o CPD na saída do reator atinge um valor máximo (mínima conversão), voltando a diminuir logo em seguida. Para temperaturas maiores que 62°C, o teor de CPD na saída sempre aumenta com o aumento de CPD na carga.

Os menores valores de CPD na saída são obtidos para baixos valores de CPD na carga e valores de temperatura acima de aproximadamente 64°C.

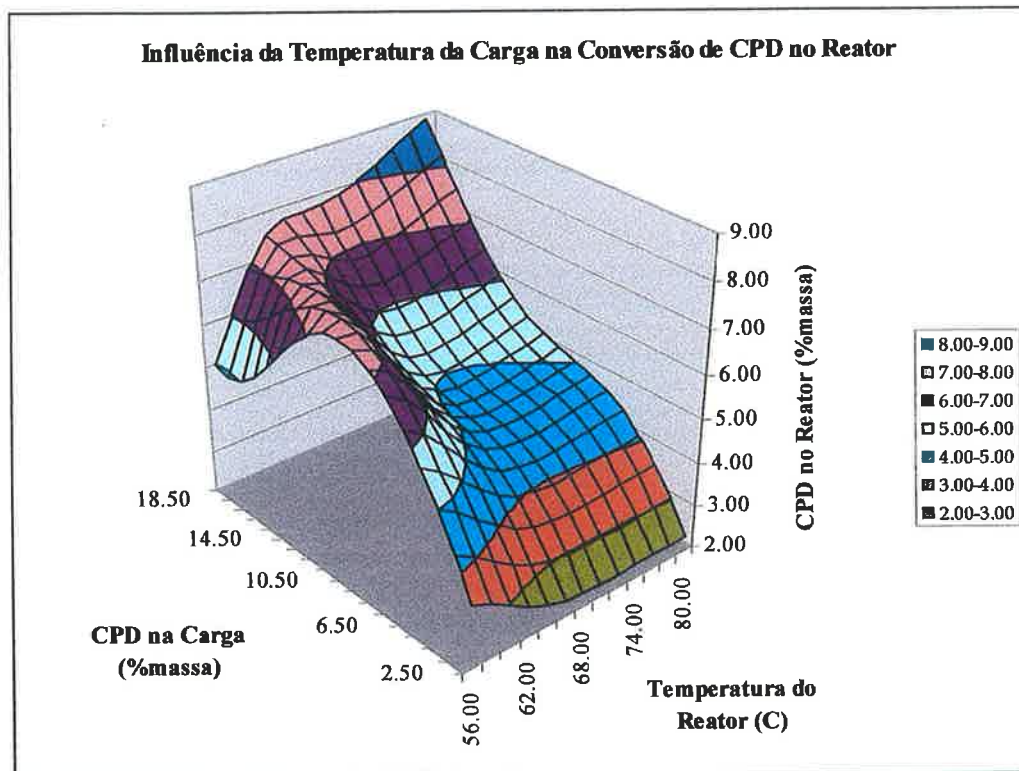


Fig. 4.45. – Influência da Temperatura da Carga do Reator na Conversão de CPD
Caso 4 - Vazão de Carga: 14,02 e %IP na Carga: 20,00

A Fig. 4.46 apresentam os resultados para a perda de IP no reator.

Observa-se que para um mesmo valor de CPD na carga, à medida que a temperatura aumenta, a perda de IP também aumenta, embora para valores inferiores a 7,50% e temperatura aproximadamente igual a 76°C, apresente um máximo, voltando a diminuir em seguida. Para valores de CPD na carga acima de 7,50, a perda de IP sempre aumenta com o aumento da temperatura.

Para um mesmo valor de temperatura, a perda de IP diminui até um valor mínimo (existe uma faixa de composição e temperatura correspondente a este ponto mínimo – região amarela da Fig. 4.46). No entanto, quanto maior a temperatura, maior o teor de CPD na carga correspondente ao ponto mínimo.

Em todos os casos, o valor da perda de IP encontra-se dentro dos limites desejáveis para o processo e operação da planta.

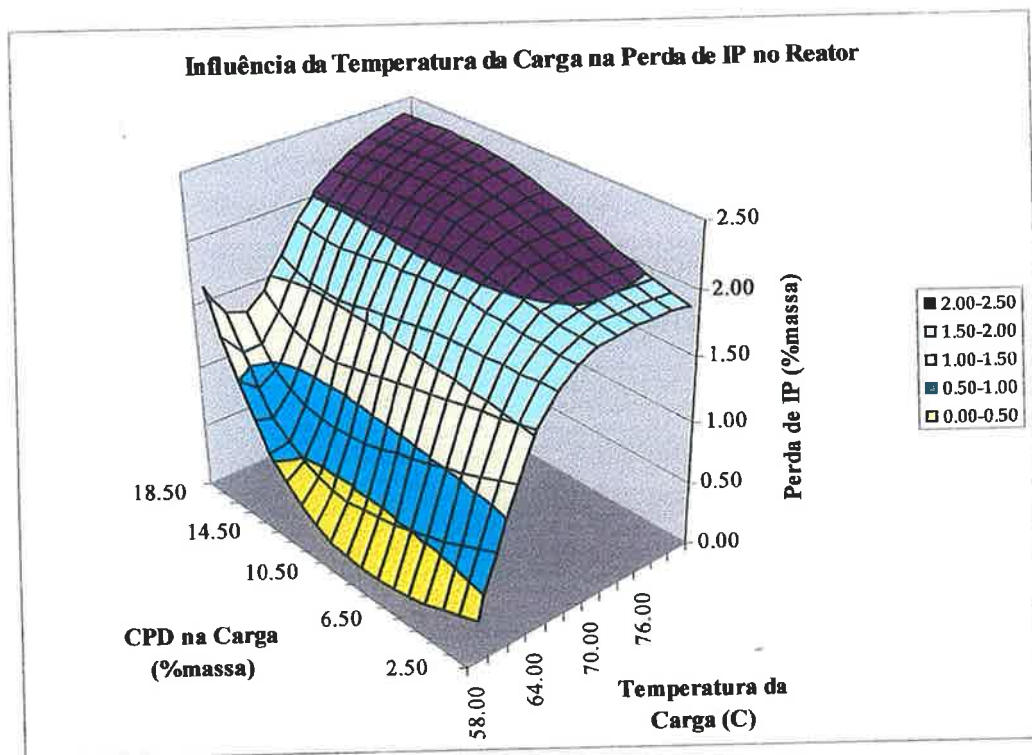


Fig. 4.46. – Influência da Temperatura da Carga do Reator na perda de IP no Reator
Caso 4 - Vazão de Carga: 14,02 e %IP na Carga: 20,00

As tabelas com os resultados da análise de sensibilidade para os casos mostrados encontram-se no ANEXO D.

Observa-se que a condição de temperatura ideal para a corrente de alimentação do reator apresenta grande variabilidade em função da grande variação em sua composição, o que acarreta comportamento distintos, mesmo diante de pequenas mudanças nas condições operacionais do reator. Uma outra explicação para esta alta sensibilidade às condições operacionais pode ser a alta complexidade do mecanismo das reações que ocorrem no processo.

Um fator de grande importância neste tipo de análise é a qualidade dos dados usados para treinamento da rede neural. Qualquer distorção na informação é representada pelo modelo gerado. Neste caso, foi observado, durante a análise e tratamento dos dados, que a temperatura do reator necessita um controle mais rigoroso. O objetivo da reação é a redução do teor de CPD da carga da unidade,

de modo a especificar a corrente de alimentação da seção de extração. Assim, cargas com baixos teores de CPD não necessitam passar por este processo. No entanto, verificam-se, na prática, longos períodos com temperaturas elevadas em condições para as quais a reação não precisaria ocorrer ou seria suficiente apenas uma temperatura mais baixa. Por outro lado, verifica-se também a ocorrência de situações inversas.

Deve-se também considerar o fato da reação de dimerização do CPD ser exotérmica, devendo existir uma condição na qual a quantidade de calor liberado pela reação é suficiente para garantir a conversão do CPD, sem necessidade de incrementos de temperatura (Pellicci e Carneiro, 1996)

É importante ressaltar que, em todos os casos, os pontos extremos dos diagramas analisados, por serem regiões com menor número de informações que as demais, podem apresentar pequenas distorções em relação à boa representatividade do sistema.

Um outro ponto importante a ser considerado é a escolha da região operacional mais adequada como aquela que apresenta menor sensibilidade às variações das condições de processo.

Outra variável importante a ser considerada é a vazão de carga da unidade. Das análises realizadas para verificar a influência da vazão da carga sobre a temperatura de entrada no reator e, conseqüentemente, no desempenho do reator, não foi possível chegar a uma conclusão clara da existência de uma relação direta entre estas variáveis. Pode-se apenas afirmar que a variável de maior importância é o teor de CPD na carga, até mesmo porque a reação principal que ocorre no processo é dimerização do CPD.

Foi observado, que para altos valores de vazão e concentração de IP na carga, a especificação de CPD requerida na saída do reator é atingida quando o teor de CPD na carga é inferior a 3,5% para uma ampla faixa de temperatura. Quando se

têm valores altos de vazão e baixas concentrações de IP, a especificação é atingida para valores de CPD na carga inferiores a 4,5%, podendo ser verificado que quanto menor o teor de CPD na carga, menor também é a temperatura da carga.

Quando se opera a baixas vazões de carga e teores baixos de IP na carga, a especificação é atingida para valores muito baixos ou muito altos de CPD na carga, e as duas situações requerem um baixo valor de temperatura.

Para baixas vazões e alto teor de IP, a especificação de CPD na saída do reator é obtida para altos teores de CPD na carga e, quanto maior este teor, menor é a temperatura necessária para atingir a especificação.

4.5.5.2. Análise de Sensibilidade da Primeira Coluna de Remoção de Pesados – COL-01

A primeira coluna de remoção de pesados (COL-01) tem por finalidade separar a corrente concentrada em isopreno de um corte mais pesado, rico em DCPD e piperilenos. Mais especificamente, a COL-01 opera de modo a separar o n-pentano dos piperilenos. A corrente de topo é a carga da seção de extração de isopreno. A corrente de fundo alimenta a COL-12, a qual separa o corte pesado em duas correntes: uma rica em piperilenos (topo) e outra em DCPD.

Assim, neste item, o objetivo é analisar as variáveis de processo da COL-01 de modo a operá-la de forma estável e otimizada, especificando a corrente de topo para atender tanto os requerimentos da carga da COL-02 quanto às necessidades da COL-12.

As variáveis selecionadas para análise de sensibilidade foram:

- ✓ Refluxo da COL-01

- ✓ Vazão e composição da carga
- ✓ Vazão de destilado

A razão de refluxo (RR) é uma das principais variáveis operacionais de uma coluna de destilação e é definida como sendo a razão entre a vazão de refluxo (L) e a vazão de destilado(D):

$$RR = L/D \qquad 4.1$$

Pode-se afirmar que quanto maior a razão de refluxo (RR), melhor será a separação. Assim, através da análise da equação da razão de refluxo, a separação é tanto melhor quanto maior a vazão de destilado e menor a vazão de destilado.

Um ponto que confunde a análise correta da teoria em relação à prática, são os efeitos do controle nas variáveis de processo. Um exemplo típico é o caso da redução do destilado para aumentar a razão de refluxo e conseqüentemente melhorar a separação, ou seja, aumentar a concentração do composto chave leve e reduzir a do chave pesado no topo¹. Na prática, o que ocorre nos primeiros instantes após uma redução do destilado é o aparecimento do componente chave leve no fundo devido ao sistema ainda não ter atingido o estado estacionário, para a nova condição operacional. Só após algum tempo, a depender da dinâmica de cada sistema é possível observar o resultado final esperado (Neto & Carneiro, 1997).

Devido à grande variabilidade na vazão e composição da carga da unidade, é difícil determinar uma condição ótima de operação para todas as possíveis combinações destas variáveis. Assim, será mostrada a influência da vazão de refluxo da COL-01 na especificação do IP, CPD e 2M2B no topo da COL-01 em relação a um conjunto específico de composição de IP e CPD na carga. Os

¹ Por definição, o composto mais leve entre os componentes que se deseja separar ou especificar é chamado de “chave leve” e, o mais pesado “chave pesado”.

resultados são apresentados na Tabela 4.26 para condições de aumento de vazão de refluxo.

Tabela 4.26.- Influência do Aumento da Vazão de Refluxo na Vazão e Composição do Topo da COL-01

	CPD baixo	CPD médio	CPD alto
IP baixo	menor %2M2B no topo	menor %2M2B no topo	menor %2M2B no topo
	menor a vazão de topo	menor a vazão de topo (variação muito pequena)	menor a vazão de topo (variação muito pequena)
	maior %IP no topo	maior %IP no topo	menor %IP no topo
	a quantidade absoluta de IP é ~cte	maior a quantidade absoluta de IP no topo	a quantidade absoluta de IP é ~cte
	menor %CPD no topo	maior %CPD no topo	maior %CPD no topo aumenta
	% CPD dentro dos limites especificados	%CPD dentro dos limites especificados	%CPD dentro dos limites especificados
IP médio	menor %2M2B no topo	menor %2M2B no topo **	menor %2M2B no topo
	menor a vazão de topo	a vazão de topo varia muito pouco***	maior a vazão de topo
	maior %IP no topo	maior %IP no topo	maior %IP no topo
	maior a quantidade absoluta de IP no topo	maior a quantidade absoluta de IP no topo	maior a quantidade absoluta de IP no topo
	maior %CPD no topo	maior %CPD no topo	maior %CPD no topo
	%CPD muito próximo do limite inferior do intervalo do domínio	%CPD dentro dos limites especificados	%CPD dentro dos limites especificados
IP alto	menor %2M2B no topo	menor %2M2B no topo	menor %2M2B no topo
	menor a vazão de topo	menor a vazão de topo	menor a vazão de topo (variação muito pequena)
	maior %IP no topo	maior %IP no topo	menor %IP no topo
	a quantidade absoluta de IP aumenta até +/- refluxo = 40 e então diminui (variação não apreciável)	maior a quantidade absoluta de IP no topo (variação não apreciável)	maior a quantidade absoluta de IP no topo
	menor %CPD no topo	menor %CPD no topo	maior %CPD no topo (variação muito pequena)
	%CPD dentro dos limites especificados	%CPD dentro dos limites especificados	%CPD dentro dos limites especificados

Notas: ** Para baixos valores de refluxo, o teor de 2M2B cai fora do domínio da variável, até pouco abaixo do ponto correspondente à vazão de topo máxima.

*** A vazão de topo aumenta até um máximo em torno 46,00 (refluxo) e depois volta a cair. A razão de refluxo (RR), porém aumenta

A maior influência do refluxo da COL-01, na corrente rica em piperilenos (topo da COL-12), está relacionada com o teor de 2M2B no topo. Quanto maior a vazão de refluxo, menor o teor de 2M2B no topo da COL-01 (carga da seção de extração) e, conseqüentemente, maior o teor de 2M2B no topo da COL-12. Ambas são condições altamente desejáveis.

De modo geral, o aumento do refluxo da COL-01 melhora a separação, verificando-se um decréscimo da vazão de destilado, assim uma quantidade maior dos piperilenos e 2M2B pode ser recuperada no topo da COL-12. Além disso, praticamente, não ocorre perda de isopreno pelo fundo da COL-01, uma vez que este componente é mais leve que o chave leve considerado na separação (n-pentano).

Pode-se efetuar estudos de análise de sensibilidade para quase todas as variáveis de processo modeladas neste estudo. Os casos aqui apresentados são apenas alguns exemplos do que pode ser feito de modo a demonstrar as potencialidades do uso de modelos de redes neurais nesta área.

4.5.6. Otimização

No caso em estudo, o objetivo principal é a modelagem do processo através de redes neurais empregando dados históricos da planta industrial, seguida da sua otimização, empregando o modelo gerado para efetuar uma busca detalhada da região de interesse através do mapeamento completo da função objetivo no domínio das variáveis. Desta forma, um “grid” pode ser construído em tempo razoável, desde que não existam muitas variáveis sendo otimizadas. Como benefício adicional, o mapeamento completo da função objetivo permite identificar facilmente múltiplos pontos ótimos. Uma vez que o mapa é obtido, fica fácil escolher as soluções que satisfazem as restrições e dentre estas, as que apresentam melhor viabilidade econômica.

Como explicado anteriormente no item 4.33, de modo geral, uma seqüência de etapas para a resolução de problemas de otimização é seguida. Cada uma destas etapas será detalhada a seguir:

a) Análise do processo:

A análise do processo, de forma a definir as variáveis e características específicas de interesse, bem como seus limites comportamentais, foi apresentada e discutida em detalhes previamente nos itens 4.4, 4.5.1, 4.5.2 e 4.5.5.

b) Desenvolvimento do modelo do processo:

O modelo representativo do processo, identificando as variáveis dependentes e independentes, foi desenvolvido e descrito em detalhes nos itens 4.5.2 e 4.5.3.

As restrições de qualidade do produto e de segurança, impostas no procedimento de otimização, bem como as variáveis manipuladas são mostradas na Tabela 4.27. As unidades das variáveis são arbitrárias.

Tabela 4. 27 - Variáveis Manipuladas e Restrições para a otimização do processo

Notação	Descrição da Variável	
Variáveis Manipuladas		
x1(1)	Vazão de Carga da Unidade	
x1(4)	Temperatura da Carga do Reator	
x2(5)	Vazão de Refluxo da COL-01	
x2(6)	Vazão de Vapor para Refervedor da COL-01	
x3(2)	Vazão de Refluxo da COL-12	
x3(4)	Vazão de Vapor para Refervedor da COL-12	
x4(6)	Vazão de Refluxo da COL-02	
x4(5)	Vazão de Solvente da COL-02	
x4(7)	Vazão de Vapor para Refervedor da COL-02	
x5(5)	Vazão de Solvente da COL-03	
x5(6)	Vazão de Refluxo da COL-03	
x5(7)	Vazão de Vapor para Refervedor da COL-03	
x7(5)	Vazão de Água para COL-07	
x7(6)	Vazão de Água para COL-06	
x8(3)	Vazão de Refluxo da COL-10	
x8(4)	Vazão de Vapor para Refervedor da COL-10	
x9(3)	Vazão de Refluxo da COL-08	
x9(4)	Vazão de Solvente para Refervedor da COL-08	
x9(5)	Temperatura de Entrada do Solvente para Refervedor da COL-08	
x9(6)	Temperatura de Saída do Solvente para Refervedor da COL-08	
x10(3)	Vazão de Refluxo da COL-09	
Restrições		
x1(3)-y1(2)	Teor de CPD na saída do R-01	< 4
x1(2)-y1(1)	Perda de IP no R-01	< 4
y2(2)	Teor de IP no topo da COL-01	> 23
y2(3)	Teor de CPD no topo da COL-01	< 4
y2(4)	Teor de 2M2B no topo da COL-01	< 2
y3(4)	Temperatura de Fundo da COL-12	< 120
y5(2)	Teor de Butino-2 no topo da COL-03	< 2
y5(3)	Teor de CPD no topo da COL-03	< 2
y8(3)	Teor de H2O no topo da COL-10	< 23

A escolha das restrições definidas na Tabela 4.27 tem como finalidade:

- ✓ Especificar o teor de CPD e 2M2B na seção de preparação de carga, de modo a garantir a especificação do produto final
- ✓ Concentrar o teor de IP na seção de preparação de carga
- ✓ Especificar a temperatura de fundo da coluna depentanizadora de modo a evitar a decomposição do DCPD
- ✓ Especificar o teor de Butino-2 e CPD no topo da Segunda Coluna de Destilação Extrativa, de modo a garantir a especificação do produto final
- ✓ Especificar o teor de Água no topo da Primeira Coluna de Recuperação de Solvente. O teor de água no solvente circulante é um ponto crítico da unidade.

Outras restrições poderiam ser consideradas, dependendo do objetivo principal a ser atingido. Os casos aqui apresentados são apenas alguns exemplos das possibilidades da capacidade de otimização do modelo gerado.

Cada variável de entrada (manipulada) tem a sua faixa de domínio dividida em igual número de intervalos, assim, as variáveis manipuladas selecionadas, em número de 21, levam a um número de milhões de casos gerados, quando do mapeamento do “grid” das condições de processo, durante a otimização usando o modelo da rede neural construído, podendo causar problemas de dimensionalidade.

c) Determinação da função objetivo específica e o critério de otimização

O procedimento empregado na otimização não requer necessariamente uma função objetiva formal, podendo-se obter tanto otimizações qualitativas (a partir de conjunto de restrições de processo e de análise de ganhos econômicos) como quantitativas (com função objetiva econômica).

Deve-se ter em mente, nesta etapa, qual o principal objetivo a ser atendido, por exemplo, maior produção, dada especificação de produto. Assim, a função

objetivo escolhida pode atender apenas aspectos operacionais, como por exemplo, especificar quais modificações podem ser efetuadas em variáveis como vazões de reciclo, vapor para refulvedor, temperaturas, para determinados valores de vazão e composição de carga.

No caso em estudo, o principal objetivo é estabelecer condições operacionais ótimas que levem a maior produção de Isopreno dentro das especificações exigidas pelo mercado.

d) Definição matemática do processo

Esta etapa envolve a escolha de uma técnica de otimização adequada para definir matematicamente o processo em estudo. Como o problema trata da otimização não linear com restrições e o interesse é demonstrar a capacidade do modelo em fornecer algumas condições operacionais ótimas dentro de alguns limites factíveis, a opção é pelo método de busca direta.

Durante o procedimento de otimização do modelo global, optou-se por dividir o sistema em três partes, de acordo com a divisão já previamente estabelecida, isto é: 1- Seção de Preparação de Carga; 2- Destilação extrativa e Regeneração de Solvente; 3- Fracionamento. Esta abordagem foi aplicada devido à possibilidade de surgimento de problemas de dimensionalidade face ao grande número de variáveis a ser otimizadas (21). Com a divisão em etapas, este problema é eliminado, sendo, então, gerado um número menor de casos, quando do mapeamento do “grid” das condições de processo. A partir deste mapeamento, as condições mais convenientes são selecionadas. Utilizou-se um programa desenvolvido, no Laboratório de Simulação e Controle de Processos (LSCP) do Departamento de Engenharia Química da Universidade de São Paulo, por Nascimento e Alves (2003).

Assim, primeiramente foi feita a otimização da seção de preparação de carga (REDE 1, REDE 2 e REDE 3), seguida da otimização da seção de destilação extrativa e regeneração de solvente (REDE 4, REDE 5, REDE 6, REDE 7, REDE 8) e por último da seção de fracionamento (REDE 9 e REDE 10). Os valores ótimos gerados em cada etapa são utilizados como dados de entrada para otimização da etapa seguinte. O procedimento de otimização de cada etapa ocorre como descrito previamente, isto é, o modelo gerado é empregado para efetuar uma busca detalhada da região de interesse através do mapeamento completo da função objetivo no domínio das variáveis.

Como exemplo para otimização, foi considerado o Caso 2 apresentado no item 4.54 e Tabelas C.1 e C.2 (ANEXO C). A Tabela 4.28 mostra, uma vez mais, as variáveis de entrada para este caso.

Tabela 4.28 – Variáveis de Entrada para o Caso Base para Otimização

x1(1)	13,99	x5(1)	64,67	x8(1)	4,39
x1(2)	20,72	x5(2)	88,40	x8(2)	78,26
x1(3)	14,90	x5(3)	1,00	x8(3)	0,79
x1(4)	59,80	x5(4)	90,89	x8(4)	1,63
x2(1)	13,50	x5(5)	14,02	x9(1)	2,67
x2(2)	62,61	x5(6)	15,01	x9(2)	29,05
x2(3)	17,33	x5(7)	5,54	x9(3)	13,05
x2(4)	4,79	x6(1)	63,34	x9(4)	53,92
x2(5)	44,24	x6(2)	28,48	x9(5)	102,28
x2(6)	9,40	x6(3)	6,71	x9(6)	70,68
x3(1)	4,89	x6(4)	17,65	x9(1)	2,67
x3(2)	2,80	x6(5)	47,83	x9(2)	29,05
x3(3)	65,64	x6(6)	1,18	x9(3)	13,05
x3(4)	1,08	x6(7)	92,33	x9(4)	53,92
x4(1)	8,24	x7(1)	7,35	x9(5)	102,28
x4(2)	24,37	x7(2)	2,55	x9(6)	70,68
x4(3)	2,87	x7(3)	1,46	x10(1)	61,58
x4(4)	1,06	x7(4)	1,38	x10(2)	2,34
x4(5)	69,91	x7(5)	2,45	x10(3)	13,22
x4(6)	12,75	x7(6)	2,50	x10(4)	13,05
x4(7)	3,98			x10(5)	18,16
				x10(6)	64,19

A primeira etapa foi fornecer a especificação da corrente de carga do reator, isto é, vazão total e concentração de Isopreno (IP) e Ciclopentadieno (CPD). Então, o mapeamento da solução foi iniciado mudando-se os valores das variáveis de processo manipuladas e, nesta etapa, apenas as soluções que satisfizeram as condições operacionais requeridas foram aceitas. Estas condições são as restrições de concentração de CPD e perda de IP no reator, concentrações de IP, CPD e 2M2B no produto de topo da primeira coluna de destilação e temperatura do fundo da coluna de pentanizadora, como mostradas na Tabela 4.27.

A etapa seguinte foi escolher qualitativamente as condições operacionais que poderiam minimizar os custos de energia; estas condições são a menor vazão de refluxo, a menor vazão de vapor para o refeedor e a menor temperatura de carga do reator.

Deve-se ter em mente a grande variabilidade das condições da carga da unidade, o que leva a diferentes condições ótimas de operação, como pôde ser observado no item 4.5.5, referente à análise de sensibilidade das variáveis de processo.

A faixa de cada variável de entrada foi dividida em quatro intervalos, assim, cerca de 4100 casos foram calculados usando o modelo da rede neural construída, destes apenas 255 satisfizeram as restrições especificadas, a partir dos quais, os mais convenientes foram selecionados de modo a atender as condições de menor custo energético (otimização qualitativa). Estes resultados são apresentados na Tabela E.1 no ANEXO E. São mostrados os resultados da simulação e os 5 casos selecionados como ótimos para a seção de preparação de carga.

Cada um destes casos gerou uma condição operacional que serviu de base para casos de otimização do segundo bloco, isto é, a seção de destilação extrativa e regeneração de solvente (REDE 4, REDE 5, REDE 6, REDE 7 e REDE 8).

A faixa de cada variável de entrada foi dividida em três intervalos, assim, cerca de 60000 casos foram calculados usando o modelo da rede neural construída para cada caso. As Tabelas E.2 a E.6 do ANEXO E apresentam os resultados selecionados de modo a atender as condições de menor custo energético (otimização qualitativa) para cada caso da Tabela E.1. O número de casos que satisfazem as restrições especificadas também pode ser visto nas tabelas.

A etapa seguinte foi escolher qualitativamente as condições operacionais que poderiam minimizar os custos de energia; estas condições são a menor vazão de refluxo, a menor vazão de vapor para o refeedor, menores temperaturas. Foram consideradas também as condições de vazão que maximizam a produção de isopreno.

Cada um dos casos considerados ótimos gerou uma condição operacional que serviu de base para casos de otimização do terceiro bloco, isto é, a seção de fracionamento (REDE 9 e REDE 10). Alguns casos do segundo bloco foram descartados por apresentarem um valor de relação solvente/carga muito acima do valor considerado ideal para o processo. Isto foi verificado para os casos de menor vazão de carga da unidade (Casos 1 e 2). Dos demais casos ótimos para os blocos 1 e 2, foram selecionados algumas condições para otimização da seção de fracionamento.

A faixa de cada variável de entrada foi dividida em quatro intervalos, assim, cerca de 1024 casos foram calculados usando o modelo da rede neural construída para cada condição operacional ótima considerada. Destes apenas 256 satisfizeram as restrições especificadas para cada caso, a partir dos quais, os mais convenientes foram selecionados de modo a atender as condições de menor custo energético (otimização qualitativa). Os resultados podem ser vistos na Tabela E.7 no ANEXO E.

Um conjunto de condições operacionais ótimas que satisfazem o critério de uma operação a um custo mais baixo para o processo global é apresentado nas Tabelas 4.29 a 4.32.

Tabela 4.29 – Otimização Global – Condição 1

Seção	Variável		Simulação	Otimização	Seção	Variável		Simulação	Otimização	Seção	Variável		Simulação	Otimização		
	Notação	Descrição				Caso 3	Notação				Descrição	Caso 3b		Notação	Descrição	3b-1
P R E P A	x1(1)	Vazão Carga	13,99	12,01	D E S T I T J L A Ç O	x4(1)	Vazão carga	7,88	7,49	F R A C I O N A M E N T O	x9(1)	Vazão carga	2,27	2,07	2,07	2,07
	x1(2)	IP carga	20,72	20,72		x4(2)	IP Carga	30,63	30,83		x9(2)	Temp carga	29,23	30,60	30,60	30,60
	x1(3)	CPD carga	14,90	14,90		x4(3)	CPD Carga	3,34	2,00		x9(3)	Refluxo	13,05	9,36	12,75	9,36
	x1(4)	Temp Carga	59,80	67,99		x4(4)	2M2B Carga	0,59	0,96		x9(4)	Vazão solv. referv.	53,92	38,26	51,12	38,26
	y1(1)	IP reator	20,09	20,07		x4(5)	Solvente	56,10	57,50		x9(5)	Temp entra solv referv.	102,28	102,05	103,74	103,18
	y1(2)	D CPD	8,17	12,41		x4(6)	Refluxo	12,75	10,62		x9(6)	Temp sai solv referv.	70,68	70,68	70,68	70,68
	x2(1)	Vazão Carga	13,99	12,01		x4(7)	Vapor Referv	3,98	2,40		y9(1)	Temp vaso	39,98	40,14	41,99	40,26
	x2(2)	Temp Carga	62,61	62,61		y4(1)	Vazão topo I	3,53	3,76		y9(2)	Temp fundo	62,09	63,62	61,72	62,17
	x2(3)	IP carga	20,09	20,07		y4(2)	Vazão topo II	3,49	4,14		y9(3)	Vazão fundo	1,90	2,64	1,78	2,57
	x2(4)	CPD carga	6,73	2,49		y4(3)	n-C5 topo	29,77	29,69		x10(1)	Temp carga	62,09	63,62	61,72	62,17
x2(5)	Refluxo	44,24	31,22	y4(4)	n-C5 topo	38,81	41,14	x10(2)	Vazão carga	1,90	2,64	1,78	2,57			
x2(6)	Vapor Referv	9,40	7,94	y4(5)	Vazão fundo	53,53	53,76	x10(3)	Refluxo	13,22	14,10	14,10	14,39			
y2(1)	Vazão Topo	7,88	7,49	y4(6)	Temp Fundo	85,96	84,42	x10(4)	Vazão cond. Referv.	13,05	13,05	13,05	13,05			
y2(2)	IP topo	30,63	30,83	x5(1)	Vazão Carga I	53,53	53,76	x10(5)	Vazão solv. Referv.	18,16	18,16	18,16	18,16			
y2(3)	CPD topo	3,34	2,00	x5(2)	Temp Carga I	85,96	84,42	x10(6)	Temp sai solv referv	64,19	64,19	64,19	64,19			
y2(4)	2M2B topo	0,59	0,96	x5(3)	Vazão Carga II	1,18	1,00	y10(1)	Vazão topo	1,87	2,48	1,91	2,04			
y2(5)	Vazão Fundo	5,63	4,60	x5(4)	Temp Carga II	92,33	90,89	y10(2)	CPD topo	0,37	0,15	0,17	0,16			
x3(1)	Vazão Carga	5,63	4,60	x5(5)	Solvente	14,02	13,00	y10(3)	Temp fundo	52,61	53,27	53,43	51,60			
x3(2)	Refluxo	2,80	2,05	x5(6)	Refluxo	15,01	14,74									
x3(3)	Temp Carga	65,64	65,64	x5(7)	Vapor Referv	5,54	5,74									
x3(4)	Vapor Referv	1,08	0,83	y5(1)	Vazão Topo	2,51	2,21									
y3(1)	Vazão Topo	2,81	2,54	y5(2)	Butino2 topo	1,43	1,67									
y3(2)	Temp topo	51,11	51,00	y5(3)	CPD topo	4,12	1,39									
y3(3)	Vazão Fundo	2,75	2,01	y5(4)	Retirada Latera	1,72	1,67									
y3(4)	Temp fundo	125,96	118,45	y5(5)	Temp # 55	98,70	99,36									
				y5(6)	Vazão Fundo	68,25	67,81									
				y5(7)	Temp Fundo	102,60	103,07									
				y5(8)	H2O no solv.	14,71	14,49									
				x6(1)	Temp Topo	63,34	63,34									
				x6(2)	Temp Vaso	28,48	28,48									
				x6(3)	IP Topo	6,71	6,71									
				x6(4)	Butino-2 topo	17,65	17,65									
				x6(5)	CPD Topo	47,83	47,83									
				x6(6)	Vazão Fundo	1,18	1,00									
				x6(7)	Temp Fundo	92,33	90,89									
				y6(1)	Vazão Carga	1,77	1,57									
				y6(2)	Refluxo	1,50	1,40									
				x7(1)	Vazão carga I	7,82	8,67									
				x7(2)	Vazão carga II	2,51	2,21									
				x7(3)	Butino2 carga	1,43	1,67									
				x7(4)	CPD carga	4,12	1,39									
				x7(5)	Vazão água I	2,45	2,29									
				x7(6)	Vazão água II	2,50	2,48									
				y7(1)	Vazão topo	2,27	2,07									
				y7(2)	Temp topo	29,23	30,59									
				y7(3)	Vazão fundo	6,15	5,99									
				x8(1)	Vazão carga	6,15	5,99									
				x8(2)	Temp carga	78,26	78,26									
				x8(3)	Refluxo	0,79	1,05									
				x8(4)	Vapor Referv	1,63	1,26									
				y8(1)	Vazão topo	1,67	1,29									
				y8(2)	Temp topo	100,63	100,63									
				y8(3)	H2O topo	20,30	20,24									
				y8(4)	Temp # 10	105,45	105,50									

Tabela 4.30. – Otimização Global – Condição 2

Seção	Variável		Simulação	Otimização	Seção	Variável		Simulação	Otimização	Seção	Variável		Simulação	Otimização		
	Notação	Descrição				Caso 4	Notação				Descrição	Caso 4b		Notação	Descrição	4b-1
P R E P A R A Ç Ã O	x1(1)	Vazão Carga	13,99	12,01	D E S T A C I O	x4(1)	Vazão carga	7,88	7,17	F R A C T O R I O	x9(1)	Vazão carga	2,27	2,28	2,28	2,28
	x1(2)	IP carga	20,72	20,72		x4(2)	IP Carga	30,63	29,71		x9(2)	Temp carga	29,23	29,31	29,31	29,31
	x1(3)	CPD carga	14,90	14,90		x4(3)	CPD Carga	3,34	2,07		x9(3)	Refluxo	13,05	9,36	12,75	9,36
	x1(4)	Temp Carga	59,80	67,99		x4(4)	2M2B Carga	0,59	0,88		x9(4)	Vazão solv. referv.	53,92	38,26	51,12	38,26
	y1(1)	IP reator	20,09	20,07		x4(5)	Solvente	56,10	56,89		x9(5)	Temp entra solv referv	102,28	102,05	102,61	103,18
	y1(2)	D CPD	8,17	12,41		x4(6)	Refluxo	12,75	15,52		x9(6)	Temp sai solv referv	70,68	70,68	70,68	70,68
	x2(1)	Vazão Carga	13,99	12,01		x4(7)	Vapor Referv	3,98	3,49		y9(1)	Temp vaso	39,98	39,40	41,57	40,53
	x2(2)	Temp Carga	62,61	62,61		y4(1)	Vazão topo I	3,53	3,06		y9(2)	Temp fundo	62,09	63,89	61,37	62,66
	x2(3)	IP carga	20,09	26,07		y4(2)	Vazão topo II	3,49	3,44		y9(3)	Vazão fundo	1,90	2,69	1,88	2,64
	x2(4)	CPD carga	6,73	2,49		y4(3)	i-C5 topo	29,77	28,00		x10(1)	Temp carga	62,09	63,89	61,37	62,66
x2(5)	Refluxo	44,24	30,72	y4(4)	n-C5 topo	38,81	37,53	x10(2)	Vazão carga	1,90	2,69	1,88	2,64			
x2(6)	Vapor Referv	9,40	7,94	y4(5)	Vazão fundo	53,53	55,16	x10(3)	Refluxo	13,22	14,10	14,10	14,39			
y2(1)	Vazão Topo	7,88	7,17	y4(6)	Temp Fundo	85,96	86,73	x10(4)	Vazão cond. Referv.	13,05	13,05	13,05	13,05			
y2(2)	IP topo	30,63	29,71	x5(1)	Vazão Carga I	53,53	55,16	x10(5)	Vazão solv. Referv.	18,16	18,16	18,16	18,16			
y2(3)	CPD topo	3,34	2,07	x5(2)	Temp Carga I	85,96	86,73	x10(6)	Temp sai solv referv	64,19	64,19	64,19	64,19			
y2(4)	2M2B topo	0,59	0,88	x5(3)	Vazão Carga II	1,18	1,00	y10(1)	Vazão topo	1,87	2,49	1,97	2,37			
y2(5)	Vazão Fundo	5,63	4,77	x5(4)	Temp Carga II	92,33	90,89	y10(2)	CPD topo	0,37	0,11	0,14	0,14			
x3(1)	Vazão Carga	5,63	4,77	A	x5(5)	Solvente	14,02	14,30	y10(3)	Temp fundo	52,61	53,35	53,06	52,40		
x3(2)	Refluxo	2,80	1,60	T	x5(6)	Refluxo	15,01	14,74								
x3(3)	Temp Carga	65,64	65,64	I	x5(7)	Vapor Referv	5,54	4,82								
x3(4)	Vapor Referv	1,08	0,83	V	y5(1)	Vazão Topo	2,51	2,51								
y3(1)	Vazão Topo	2,81	2,34	A	y5(2)	Butino2 topo	1,43	1,62								
y3(2)	Temp topo	51,11	51,67		y5(3)	CPD topo	4,12	1,93								
y3(3)	Vazão Fundo	2,75	2,36	E	y5(4)	Retirada Latera	1,72	1,53								
y3(4)	Temp fundo	125,96	119,75		y5(5)	Temp # 55	98,70	98,04								
					y5(6)	Vazão Fundo	68,25	68,69								
					y5(7)	Temp Fundo	102,60	101,70								
					y5(8)	H2O no solv.	14,71	16,18								
					x6(1)	Temp Topo	63,34	63,34								
					x6(2)	Temp Vaso	28,48	28,48								
					x6(3)	IP Topo	6,71	6,71								
					x6(4)	Butino-2 topo	17,65	17,65								
					x6(5)	CPD Topo	47,83	47,83								
					x6(6)	Vazão Fundo	1,18	1,00								
					x6(7)	Temp Fundo	92,33	90,89								
					y6(1)	Vazão Carga	1,77	1,57								
					y6(2)	Refluxo	1,50	1,40								
					x7(1)	Vazão carga I	7,82	7,27								
					x7(2)	Vazão carga II	2,51	2,51								
					x7(3)	Butino2 carga	1,43	1,62								
					x7(4)	CPD carga	4,12	1,93								
					x7(5)	Vazão água I	2,45	2,29								
					x7(6)	Vazão água II	2,50	2,48								
					y7(1)	Vazão topo	2,27	2,28								
					y7(2)	Temp topo	29,23	29,31								
					y7(3)	Vazão fundo	6,15	5,93								
					x8(1)	Vazão carga	6,15	5,93								
					x8(2)	Temp carga	78,26	78,26								
					x8(3)	Refluxo	0,79	0,60								
					x8(4)	Vapor Referv	1,63	1,26								
					y8(1)	Vazão topo	1,67	1,09								
					y8(2)	Temp topo	100,63	100,63								
					y8(3)	H2O topo	20,30	20,37								
					y8(4)	Temp # 10	105,45	105,43								

Tabela 4.31. – Otimização Global – Condição 3

Seção	Variável		Simulação	Otimização	Seção	Variável		Simulação	Otimização	Seção	Variável		Simulação	Otimização		
	Notação	Descrição				Caso 4	Notação				Descrição	Caso 4d		Notação	Descrição	4d-1
P R E P A A Ç Ã O	x1(1)	Vazão Carga	13,99	12,01	D E S T I L A A Ç Ã O	x4(1)	Vazão carga	7,88	7,17	F R A C I O N A M E N T O	x9(1)	Vazão carga	2,27	2,28	2,28	2,28
	x1(2)	IP carga	20,72	20,72		x4(2)	IP Carga	30,63	29,71		x9(2)	Temp carga	29,23	29,56	29,56	29,56
	x1(3)	CPD carga	14,90	14,90		x4(3)	CPD Carga	3,34	2,07		x9(3)	Refluxo	13,05	9,36	12,75	9,36
	x1(4)	Temp Carga	59,80	67,99		x4(4)	2M2B Carga	0,59	0,88		x9(4)	Vazão solv. referv.	53,92	38,26	51,12	38,26
	y1(1)	IP reator	20,09	20,07		x4(5)	Solvente	56,10	56,89		x9(5)	Temp entra solv referv	102,28	102,61	102,61	103,18
	y1(2)	D CPD	8,17	12,41		x4(6)	Refluxo	12,75	15,32		x9(6)	Temp sai solv referv	70,68	70,68	70,68	70,68
	x2(1)	Vazão Carga	13,99	12,01		x4(7)	Vapor Referv	3,98	3,49		y9(1)	Temp vaso	39,98	40,65	41,69	40,92
	x2(2)	Temp Carga	62,61	62,61		y4(1)	Vazão topo I	3,53	3,06		y9(2)	Temp fundo	62,09	63,34	61,51	62,85
	x2(3)	IP carga	20,09	20,07		y4(2)	Vazão topo II	3,49	3,44		y9(3)	Vazão fundo	1,90	2,68	1,87	2,65
	x2(4)	CPD carga	6,73	2,49		y4(3)	i-C5 topo	29,77	28,00		x10(1)	Temp carga	62,09	63,34	61,51	62,85
x2(5)	Refluxo	44,24	36,72	y4(4)	n-C5 topo	38,81	37,53	x10(2)	Vazão carga	1,90	2,68	1,87	2,65			
x2(6)	Vapor Referv	9,40	7,94	y4(5)	Vazão fundo	53,53	55,16	x10(3)	Refluxo	13,22	14,10	14,10	11,39			
y2(1)	Vazão Topo	7,88	7,17	y4(6)	Temp Fundo	85,96	86,73	x10(4)	Vazão cond. Referv.	13,05	13,05	13,05	13,05			
y2(2)	IP topo	30,63	29,71	x5(1)	Vazão Carga I	53,53	55,16	x10(5)	Vazão solv. Referv.	18,16	18,16	18,16	18,16			
y2(3)	CPD topo	3,34	2,07	x5(2)	Temp Carga I	85,96	86,73	x10(6)	Temp sai solv referv	64,19	64,19	64,19	64,19			
y2(4)	2M2B topo	0,59	0,88	x5(3)	Vazão Carga II	1,18	1,60	y10(1)	Vazão topo	1,87	2,49	1,95	2,35			
y2(5)	Vazão Fundo	5,63	4,77	x5(4)	Temp Carga II	92,33	90,89	y10(2)	CPD topo	0,37	0,19	0,17	0,14			
x3(1)	Vazão Carga	5,63	4,77	x5(5)	Solvente	14,02	14,30	y10(3)	Temp fundo	52,61	53,26	53,21	52,56			
x3(2)	Refluxo	2,80	1,60	x5(6)	Refluxo	15,01	14,74									
x3(3)	Temp Carga	65,64	65,64	x5(7)	Vapor Referv	5,54	4,82									
x3(4)	Vapor Referv	1,08	0,83	y5(1)	Vazão Topo	2,51	2,51									
y3(1)	Vazão Topo	2,81	2,84	y5(2)	Butino2 topo	1,43	1,62									
y3(2)	Temp topo	51,11	51,87	y5(3)	CPD topo	4,12	1,93									
y3(3)	Vazão Fundo	2,75	2,36	y5(4)	Retirada Latera	1,72	1,53									
y3(4)	Temp fundo	125,96	119,78	y5(5)	Temp # 55	98,70	98,04									
				y5(6)	Vazão Fundo	68,25	68,69									
				y5(7)	Temp Fundo	102,60	101,70									
				y5(8)	H2O no solv.	14,71	16,18									
				x6(1)	Temp Topo	63,34	63,34									
				x6(2)	Temp Vaso	28,48	28,48									
				x6(3)	IP Topo	6,71	6,71									
				x6(4)	Butino-2 topo	17,65	17,65									
				x6(5)	CPD Topo	47,83	47,83									
				x6(6)	Vazão Fundo	1,18	1,60									
				x6(7)	Temp Fundo	92,33	90,89									
				y6(1)	Vazão Carga	1,77	1,57									
				y6(2)	Refluxo	1,50	1,50									
				x7(1)	Vazão carga I	7,82	7,27									
				x7(2)	Vazão carga II	2,51	2,51									
				x7(3)	Butino2 carga	1,43	1,62									
				x7(4)	CPD carga	4,12	1,93									
				x7(5)	Vazão água I	2,45	2,69									
				x7(6)	Vazão água II	2,50	2,48									
				y7(1)	Vazão topo	2,27	2,28									
				y7(2)	Temp topo	29,23	29,36									
				y7(3)	Vazão fundo	6,15	6,16									
				x8(1)	Vazão carga	6,15	6,16									
				x8(2)	Temp carga	78,26	78,26									
				x8(3)	Refluxo	0,79	0,60									
				x8(4)	Vapor Referv	1,63	1,26									
				y8(1)	Vazão topo	1,67	1,09									
				y8(2)	Temp topo	100,63	100,63									
				y8(3)	H2O topo	20,30	20,42									
				y8(4)	Temp # 10	105,45	105,43									

Tabela 4.32. – Otimização Global – Condição 4

Seção	Variável		Simulação	Otimização	Seção	Variável		Simulação	Otimização	Seção	Variável		Otimização			
	Notação	Descrição				Notação	Descrição				Notação	Descrição	5a-1	5a-2	5a-3	
P R E P A R C Ã O D E C A R G A	x1(1)	Vazão Carga	13,99	13,27	D E S T I L A L A C Ã O O E X T R A C O L U M N A R I O S	x4(1)	Vazão carga	7,88	7,46	F R A C I O N I O M E N T A D O	x9(1)	Vazão carga	2,27	2,05	2,05	2,05
	x1(2)	IP carga	20,72	20,72		x4(2)	IP Carga	30,63	32,66		x9(2)	Temp carga	29,23	30,38	30,38	30,38
	x1(3)	CPD carga	14,90	14,90		x4(3)	CPD Carga	3,34	1,93		x9(3)	Refluxo	13,05	9,36	12,75	9,36
	x1(4)	Temp Carga	59,80	74,08		x4(4)	2M2B Carga	0,59	1,10		x9(4)	Vazão solv. referv	53,92	38,26	51,12	38,26
	y1(1)	IP reator	20,09	19,21		x4(5)	Solvente	56,10	57,61		x9(5)	Temp entra solv referv	102,28	102,05	103,74	102,61
	y1(2)	D CPD	8,17	12,16		x4(6)	Refluxo	12,75	10,62		x9(6)	Temp sai solv referv	70,68	70,68	70,68	70,68
	x2(1)	Vazão Carga	13,99	13,27		x4(7)	Vapor Referv	3,98	4,59		y9(1)	Temp vaso	39,98	39,66	41,91	39,99
	x2(2)	Temp Carga	62,61	62,61		y4(1)	Vazão topo I	3,53	3,58		y9(2)	Temp fundo	62,09	61,58	61,71	62,63
	x2(3)	IP carga	20,09	19,21		y4(2)	Vazão topo II	3,49	3,62		y9(3)	Vazão fundo	1,90	2,64	1,76	2,61
	x2(4)	CPD carga	6,73	2,74		y4(3)	i-C5 topo	29,77	30,62		x10(1)	Temp carga	62,09	63,58	61,71	62,63
x2(5)	Refluxo	44,24	31,32	y4(4)	n-C5 topo	38,81	40,79	x10(2)	Vazão carga	1,90	2,64	1,76	2,61			
x2(6)	Vapor Referv	9,40	7,94	y4(5)	Vazão fundo	53,53	53,89	E	x10(3)	Refluxo	13,22	14,10	14,10	11,39		
y2(1)	Vazão Topo	7,88	7,46	y4(6)	Temp Fundo	85,96	85,09	N	x10(4)	Vazão cond. Referv.	13,05	13,05	13,05	13,05		
y2(2)	IP topo	30,63	32,65	x5(1)	Vazão Carga I	53,53	53,89	T	x10(5)	Vazão solv. Referv.	18,16	18,16	18,16	18,16		
y2(3)	CPD topo	3,34	1,92	x5(2)	Temp Carga I	85,96	85,09	O	x10(6)	Temp sai solv referv	64,19	64,19	64,19	64,19		
y2(4)	2M2B topo	0,59	1,10	x5(3)	Vazão Carga II	1,18	1,00	y10(1)	Vazão topo	1,87	2,97	1,91	2,35			
y2(5)	Vazão Fundo	5,63	5,13	x5(4)	Temp Carga II	92,33	90,89	y10(2)	CPD topo	0,37	0,16	0,17	0,14			
x3(1)	Vazão Carga	5,63	5,13	A	x5(5)	Solvente	14,02	13,00	y10(3)	Temp fundo	52,61	53,26	53,43	52,26		
x3(2)	Refluxo	2,80	1,69	T	x5(6)	Refluxo	15,01	14,74								
x3(3)	Temp Carga	65,64	65,64	I	x5(7)	Vapor Referv	5,54	5,24								
x3(4)	Vapor Referv	1,08	0,83	V	y5(1)	Vazão Topo	2,51	2,22								
y3(1)	Vazão Topo	2,81	2,29	A	y5(2)	Butino2 topo	1,43	1,65								
y3(2)	Temp topo	51,11	51,45		y5(3)	CPD topo	4,12	1,05								
y3(3)	Vazão Fundo	2,75	2,87		y5(4)	Retirada Lateral	1,72	1,66								
y3(4)	Temp fundo	125,96	116,40		y5(5)	Temp # 55	98,70	99,51								
					E	y5(6)	Vazão Fundo	68,25	67,91							
						y5(7)	Temp Fundo	102,60	103,19							
						y5(8)	H2O no solv.	14,71	14,64							
					R	x6(1)	Temp Topo	63,34	63,34							
					R	x6(2)	Temp Vaso	28,48	28,48							
					R	x6(3)	IP Topo	6,71	6,71							
					C	x6(4)	Butino-2 topo	17,65	17,65							
					U	x6(5)	CPD Topo	47,83	47,83							
					P	x6(6)	Vazão Fundo	1,18	1,60							
					E	x6(7)	Temp Fundo	92,33	90,89							
					R	y6(1)	Vazão Carga	1,77	1,57							
					A	y6(2)	Refluxo	1,50	1,40							
					Ç	x7(1)	Vazão carga I	7,82	7,97							
					Ã	x7(2)	Vazão carga II	2,51	2,22							
					O	x7(3)	Butino2 carga	1,43	1,65							
						x7(4)	CPD carga	4,12	1,05							
					D	x7(5)	Vazão água I	2,45	2,29							
					E	x7(6)	Vazão água II	2,50	2,48							
						y7(1)	Vazão topo	2,27	2,05							
					S	y7(2)	Temp topo	29,23	30,37							
					O	y7(3)	Vazão fundo	6,15	5,92							
					L	x8(1)	Vazão carga	6,15	5,92							
					V	x8(2)	Temp carga	78,26	78,26							
					E	x8(3)	Refluxo	0,79	1,05							
					N	x8(4)	Vapor Referv	1,63	1,26							
					T	y8(1)	Vazão topo	1,67	1,30							
					E	y8(2)	Temp topo	100,63	109,63							
						y8(3)	H2O topo	20,30	20,24							
						y8(4)	Temp # 10	105,45	105,51							

Apesar dos vários casos que satisfazem as restrições estabelecidas, a solução final de cada caso estudado apresenta apenas poucas condições operacionais que satisfazem o critério de uma operação a um custo mais baixo.

4.6. Conclusões

No caso estudado, a metodologia seguida para construção do modelo global, como uma seqüência de redes neurais interconectadas por ciclos e variáveis de correntes intermediárias, apresentou muito bons resultados, o que permitiu uma boa compreensão do comportamento do processo analisado e das variáveis importantes para a operação da planta. Conseqüentemente o modelo gerado é uma ferramenta valiosa para engenheiros e supervisores de processo, trazendo mais benefícios que os modelos convencionais. Além disso, foi possível, através da análise de dados detectar erros de instrumentação em certos períodos de operação e corrigi-los, quando possível ou elimina-los, baseando-se em procedimentos estatísticos.

Os resultados da simulação do modelo global foram coerentes com o comportamento da planta, evidenciando que a metodologia desenvolvida pode ser empregada com segurança para a representação do processo de produção de isopreno.

Obtendo-se um modelo global muito bem ajustado e representativo do processo, o modelo pode ser usado para simular o processo com grande rapidez. Milhões de simulações podem ser executadas em segundos, o que permite sua utilização em procedimentos de otimização de processo, controle de processo, controle “on-line”, análise de restrições de processo, análise de sensibilidade das variáveis do processo e ate mesmo como analisadores virtuais, entre outras aplicações.

Neste estudo, o modelo gerado foi utilizado para otimização do processo. Tal procedimento, baseado em um “grid” de busca da solução do problema, usando redes neurais como um modelo da planta industrial, tem a vantagem de mapear todas as soluções possíveis. Este procedimento permite definir o domínio no qual a solução ótima é localizada. Soluções quantitativas e qualitativas podem ser rapidamente analisadas. Além disso, para atingir pontos ótimos de operação, não é necessário utilizar uma função objetivo explícita, uma vez que é possível trabalhar apenas com desigualdades.

Os resultados obtidos nos estudos de análise de sensibilidade e otimização fornecem sugestões de condições operacionais possíveis de aplicação na unidade, ressaltando o aspecto prático do estudo desenvolvido.

A modelagem via redes neurais depende fortemente da qualidade dos dados fornecidos para o seu treinamento, bem como do domínio das variáveis. Assim, para a unidade de produção de isopreno estudada, alguns resultados poderiam ser melhorados se os dados empregados fossem medidos com mais rigor. Isto é, faz-se necessário uma boa manutenção dos instrumentos de medição, uma vez que foram observadas informações discrepantes em curto período de tempo, apresentando alta variabilidade. Em outros casos, foram observados valores nos limites de fundo de escala dos instrumentos por um longo período de tempo (em alguns casos, até meses). Além disso, verificou-se ainda que variáveis críticas para o processo não são analisadas quanto à sua variação ou sensibilidade em relação às diferentes condições operacionais apresentadas, o que levaria a um melhor controle destas variáveis e conseqüentemente maior estabilidade operacional poderia ser atingida.

Faz-se necessário também reavaliar a necessidade de novas medidas para variáveis importantes de modo a melhorar a confiabilidade das informações do processo e, conseqüentemente, melhorar a capacidade preditiva do modelo gerado para avaliar a planta.

5. PREDIÇÃO DE AZEÓTROPAS

5.1. Introdução

A capacidade em prever se uma dada mistura forma ou não um azeótropo, incluindo as condições requeridas para a ocorrência do azeótropo (temperatura e composição), é essencial em qualquer processo de separação. Azeótropos, geralmente, ocorrem em fracionamento e torna uma dada separação impossível por retificação simples ou podem ser utilizados para separar misturas não comumente separadas por retificação direta. Assim, o conhecimento de azeotropia é importante para o sucesso na seleção de solventes a serem utilizados em destilação azeotrópica ou extrativa ou em processos de extração.

Devido à importância crucial da azeotropia em projeto e síntese de processos de destilação, um grande número de dados experimentais sobre comportamento azeotrópico tem sido compilado ao longo dos anos (Horsley, 1973, Gmehling et al., 1994). Embora cerca de 40000 conjuntos de dados azeotrópicos tenham sido reportados, o número de possíveis azeótropos é teoricamente infinito (Kim & Simmrock, 1997). Além disso, a determinação de azeotropia apenas a partir de experimentos pode ser muito caro e consumir muito tempo. Evidentemente, a disponibilidade de métodos computacionais para a predição de azeótropos poderia não apenas reduzir este custo, mas também poupar tempo, diminuindo o espaço experimental de busca. Embora várias correlações práticas tenham sido propostas, muitas delas são empíricas, não apresentam uma base física clara e têm faixa de aplicabilidade restrita, isto é, nem sempre é possível utilizá-las para todos os tipos de misturas não-ideais sem quaisquer suposições adicionais.

A formação de misturas com ebulição constante é uma função da diferença nas temperaturas de ebulição dos componentes puros e da não idealidade da solução resultante. A não idealidade de uma mistura depende fortemente das forças intermoleculares dos componentes tais como, forças de dispersão, interações dipolo - dipolo, interações dipolo - dipolo induzido, pressão interna e pontes de hidrogênio,

dos quais o último é o efeito mais importante na formação de um azeótropo (Ewell et al., 1944; Kim & Simmrock, 1997).

Com isto em mente, o objetivo deste trabalho é utilizar informação experimental sobre azeotropia disponível para desenvolver um método alternativo para predição de formação de azeótropos entre componentes em uma mistura binária e as condições azeotrópicas (tipo, temperatura e composição), sem executar cálculos de equilíbrio líquido – vapor, isto é, sem qualquer consideração explícita da não – idealidade dos componentes da mistura na fase líquida. A metodologia proposta correlaciona dados binários azeotrópicos com propriedades macroscópicas e microscópicas dos componentes puros constituintes da mistura. A metodologia aplicada faz uso de redes neurais para correlacionar os sistemas e conseqüentemente prever o azeótropo. As propriedades macroscópicas examinadas incluem massa molecular, ponto de ebulição, constantes críticas e fator acêntrico, enquanto as propriedades moleculares consistem de parâmetros relacionados a dipolaridade molecular e a capacidade de cada componente em formar pontes de hidrogênio. A vantagem de usar redes neurais, além de sua capacidade em ajustar dados, é não ser necessário empregar modelos termodinâmicos que representem satisfatoriamente a não idealidade da fase líquida, como citado previamente.

Inicialmente, uma revisão geral sobre azeotropia será apresentada. Em seguida, a metodologia usada e a estrutura da rede serão apresentadas, incluindo a melhor seleção dos parâmetros de entrada, bem como as regras estabelecidas para discriminar se uma dada mistura binária forma ou não um azeótropo e a correspondente condição azeotrópica (tipo, temperatura e composição).

5.2. Princípios Básicos da Azeotropia

Azeótropo denomina uma mistura na qual o líquido em equilíbrio com o seu vapor têm a mesma composição. De acordo com o teorema de Gibbs-Konovalov, um azeótropo exibe um ponto de ebulição máximo ou mínimo a pressão constante, isto é,

uma pressão máxima ou mínima a temperatura constante no ponto de azeótropo (Kim & Simmrock, 1997). Este teorema é uma condição de importância prática em engenharia química, especialmente na área de síntese e projeto de processo, uma vez que limita a possibilidade de separação por destilação simples. Assim, um azeótropo é uma mistura que tem um ponto de ebulição máximo ou mínimo em relação aos pontos de ebulição das misturas com composições nas vizinhanças da sua. Os diagramas de equilíbrio de fases tipo $y-x$ mostram uma inversão na volatilidade relativa nos pontos em que ocorrem máximos de pressão ou de temperatura. No ponto de inversão a composição do vapor e a do líquido são idênticas, e a separação dos dois componentes, além deste ponto, não é possível mediante destilação convencional.

O fenômeno de azeotropia ocorre em misturas que formam soluções líquidas não ideais, cujas moléculas têm tamanho e forma diferentes e suas forças intermoleculares também diferem das forças entre moléculas semelhantes. Uma solução ideal obedece a **Lei de Raoult** ($y_i P = x_i p_i^{sat} = P_i$) significando que **a pressão parcial de qualquer composto é igual ao produto de sua pressão de vapor pela sua fração molar na fase líquida.**

O esclarecimento das grandezas fugacidade, atividade e coeficiente de atividade auxiliam no entendimento do significado da não idealidade da fase líquida relacionada com a formação de azeótropos, sendo por isso resumida nos próximos parágrafos.

A **fugacidade (f)** é uma função de estado que tem as mesmas dimensões de pressão e pode ser definida como uma função auxiliar que tem o seu valor igualado à pressão quando P tende a zero condição em que podemos considerar que uma mistura é ideal pois as moléculas estão suficientemente afastadas entre si para não gerar forças moleculares e de interações, ou seja:

$$\lim_{P \rightarrow 0} \left(\frac{f}{P} \right) = 1 \quad (5.1)$$

Em termos práticos, no entanto, a grandeza que melhor se relaciona à fugacidade de uma mistura na fase líquida, a baixas pressões, é a **atividade** do componente i (a_i) expresso por:

$$a_i = f_i / f_i^0 \quad (5.2)$$

onde f_i e f_i^0 são respectivamente a fugacidade do componente i no estado em estudo e a fugacidade do mesmo componente i no estado padrão à mesma temperatura.

A partir do conceito de atividade, define-se **coeficiente de atividade** (γ) como:

$$\gamma_i = a_i / x_i \quad (5.3)$$

Para uma **solução líquida ideal**, define-se $\gamma_i = 1$, de modo que $a_i = x_i$, ou seja, o valor da atividade tende ao valor da concentração da fase em questão.

A equação que relaciona o coeficiente de atividade γ_i e a **lei de Raoult** incluindo apenas a não idealidade da fase líquida é a seguinte:

$$P_i = P y_i = \gamma_i x_i P_i^{sat} \quad (5.4)$$

Os azeótropos classificam-se em azeótropos de mínimo e de máximo. A ocorrência de temperaturas máximas ou mínimas nos gráficos de temperatura versus composição é provocada por desvios *negativos ou positivos*, respectivamente da lei de Raoult. Estes desvios ocorrem devido a não idealidade da fase líquida nas misturas, onde para desvios negativos ($\gamma < 1$) tem-se que as forças intermoleculares são mais fortes entre os pares de moléculas distintas do que entre os pares de moléculas semelhantes quando comparado à solução ideal (forças intermoleculares iguais), o que significa que as moléculas ficam retidas mais facilmente na fase líquida. Quando ocorre desvio positivo ($\gamma > 1$) as moléculas se liberam mais facilmente para a fase vapor em relação à condição ideal, pois a atração entre as

moléculas semelhantes é maior do que entre os pares de moléculas distintas na fase líquida. A Tabela 5.1 mostra esquematicamente um resumo destas conclusões.

Tabela 5.1. - Desvios da Lei de Raoult

Gama	Desvio	Tendência	Interação
$\gamma > 1$	Positivo $P_i > P_i^{ideal}$	Moléculas diferentes mais soltas na fase líquida indo para a fase vapor	Forças intermoleculares mais fracas entre moléculas distintas em relação às semelhantes
$\gamma < 1$	Negativo $P_i < P_i^{ideal}$	Moléculas diferentes mais presas na fase líquida ficando na fase líquida	Forças intermoleculares mais fortes entre moléculas distintas em relação às semelhantes

Além disso, os azeótropos podem ser divididos em homogêneos e heterogêneos. Um azeótropo é homogêneo quando apenas uma fase líquida está presente. Um azeótropo é heterogêneo quando os desvios em relação à lei de Raoult são suficientemente grandes podendo ocorrer imiscibilidade parcial e a presença de duas fases líquidas, persistindo até a temperatura de ebulição. Nestes casos, forma-se um azeótropo de mínimo quando uma fase vapor está em equilíbrio com duas fases líquidas. O azeótropo heterogêneo constitui o mesmo obstáculo à separação por destilação simples que o azeótropo homogêneo. A miscibilidade parcial, no entanto, possibilita o processo de separação simples.

Exemplos típicos de misturas azeotrópicas são apresentados na Fig.5.1.

Até o momento, um grande número de dados experimentais tem sido obtido ao longo do tempo e a informação armazenada em bancos de dados e publicada na literatura. Apesar disto, a informação ainda é pouca frente ao número quase infinito de azeótropos teoricamente possíveis. Além disso, várias correlações independentes para predição de dados azeotrópicos têm sido propostas, no entanto, são altamente empíricas e é difícil encontrar a base física comum que pode ser aplicada genericamente a todos os tipos de misturas não ideais.

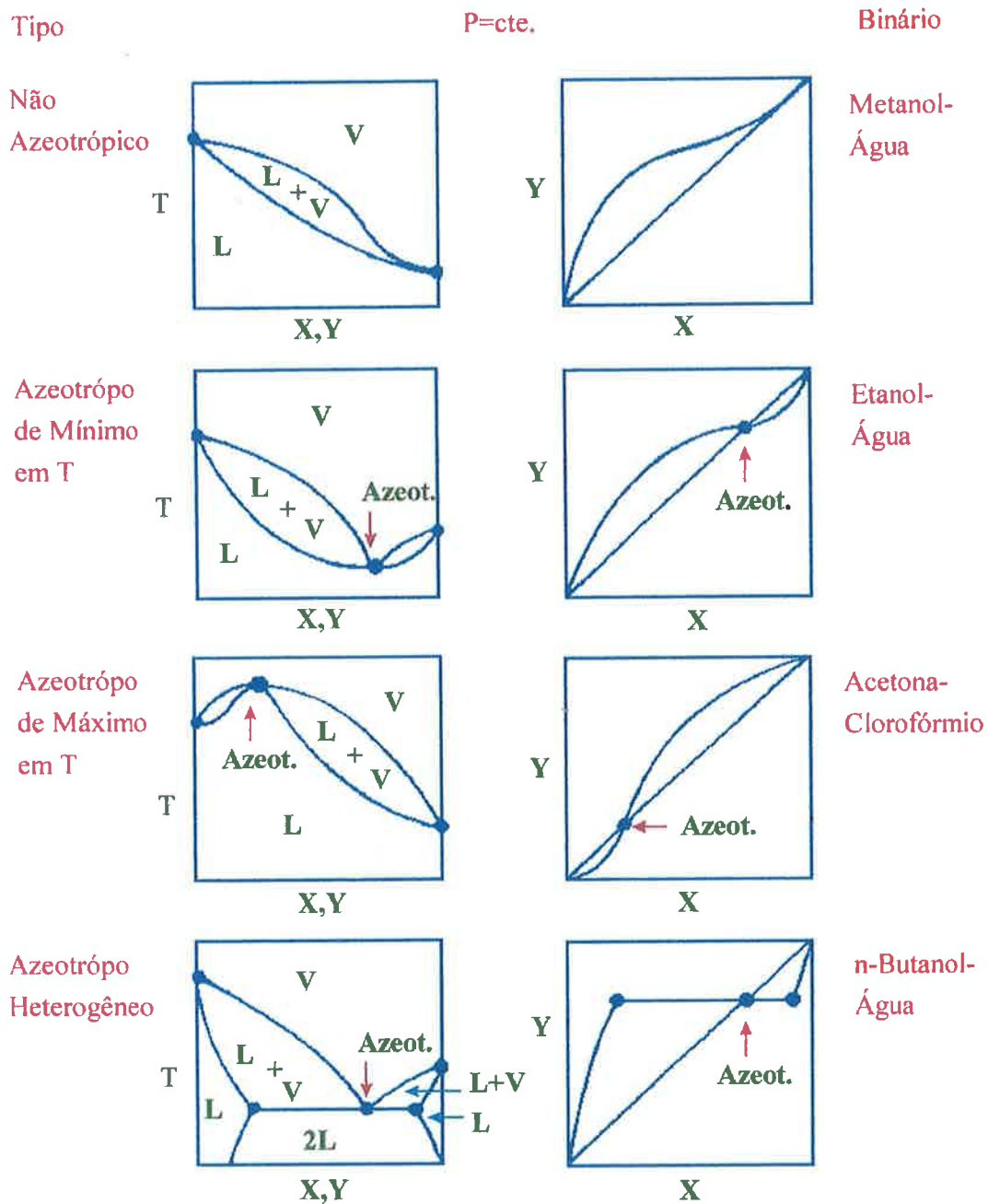


Fig.5.1 – Exemplos Típicos de Misturas Azeotrópicas

Se uma mistura é ou não azeotrópica depende essencialmente dos seguintes fatores:

- grau de não idealidade de uma mistura
- diferença entre os pontos de ebulição dos componentes puros

Em uma solução ideal, as forças intermoleculares médias da mistura não variam, assim como não há variação na energia interna ou no volume, nem calor é absorvido ou liberado quando da formação da mistura, e a pressão de vapor da mistura é uma função linear da composição molar, isto é, o sistema obedece a Lei de Raoult.

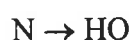
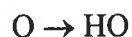
Desde que a formação de misturas azeotrópicas é uma função da não-idealidade da mistura, uma correlação dos desvios da idealidade, isto é, da Lei de Raoult, para várias misturas poderia ser de grande ajuda na predição e seleção de sistemas azeotrópicos. Como mencionado previamente, em geral, os desvios da idealidade podem ser atribuídos às contribuições de diferentes tipos de forças intermoleculares entre os componentes constituintes da mistura, tais como forças de dispersão (forças entre moléculas não polares), interações dipolo – dipolo, interações dipolo – dipolo induzido, e pontes de hidrogênio, dentre os quais este último é o efeito mais importante na formação de azeótropos.

Embora significativo progresso na azeotropia dependa basicamente do desenvolvimento da física molecular, isto é, do conhecimento suficiente das forças intermoleculares em uma mistura, e de modelos termodinâmicos que representem satisfatoriamente a fase líquida, o grau de não idealidade em misturas pode ser predita com razoável precisão com a classificação de componentes com base na capacidade de formação de pontes de hidrogênio de certos grupos.

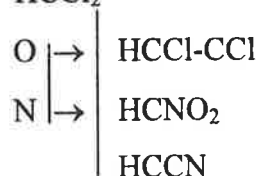
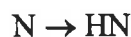
O conceito de pontes de hidrogênio é que hidrogênio pode coordenar entre duas moléculas de oxigênio, nitrogênio, ou flúor. Hidrogênio pode também coordenar entre um destes compostos doadores de átomos (oxigênio, nitrogênio, ou flúor) e um átomo de carbono, desde que um número suficiente de átomos ou grupos negativos esteja ligado ao átomo de carbono. Hidrogênio não pode coordenar entre dois átomos de carbono. A força das pontes de hidrogênio depende da natureza dos átomos entre

os quais o hidrogênio está coordenando. Elas podem ser classificadas em ligações fortes e fracas como segue:

FORTE



FRACA



Baseado nas pontes de hidrogênio, é possível classificar líquidos orgânicos em cinco classes. É, então, possível generalização, observando sistemas binários compostos de combinações destas classes. Assim, atribuindo a dois líquidos suas classes de pontes de hidrogênio apropriadas, a natureza dos desvios da idealidade e a possibilidade de formação de azeótropo podem ser preditas.

No entanto, apenas com a informação do grau de não idealidade de uma mistura, a predição de formação de azeótropos é só qualitativa. Predições mais precisas podem ser obtidas levando em conta a diferença entre os pontos de ebulição dos dois componentes puros. O efeito desta diferença na azeotropia pode ser explicado com o conceito de “faixa azeotrópica” introduzido por Swietoslawski (1950) e desenvolvido por Malensinski (1956) e Yoshimoto & Mashiko (1956). Este conceito é baseado na suposição geral que os componentes formam uma solução estritamente regular próxima do seu ponto de ebulição, a qual pode ser tratada usando o modelo quase cristalino de Guggenheim (1967), introduzindo representação explícita para a energia de interação das moléculas componentes. Uma solução estritamente regular pode ser definida como soluções com moléculas esféricas de tamanho aproximadamente iguais e com fracas energias de interação.

Quanto mais próximos os pontos de ebulição dos dois componentes, mais provável a formação de azeótropo; quanto mais ideal a solução de componentes, menos provável a formação de um sistema azeotrópico. Entretanto, para um dado par de substâncias, quanto menor o desvio da idealidade, tanto menor deve ser a diferença entre os pontos de ebulição requeridos para a formação de um azeótropo. Por outro lado, componentes com grandes diferenças nos pontos de ebulição podem não exibir um azeótropo mesmo que eles formem uma mistura com comportamento altamente não ideal.

Do ponto de vista prático, um método de predição baseado na diferença entre os pontos de ebulição dos componentes da mistura binária e sua capacidade para formar uma mistura não ideal, poderia, portanto, ser um bom candidato na correlação de azeótropos com propriedades físicas e moleculares das substâncias puras.

5.3. .. Metodologia

A metodologia proposta emprega modelos de redes neurais para a predição de sistemas binários azeotrópicos. A idéia básica é caracterizar as substâncias puras através de propriedades básicas tais como massa molecular, temperatura de ebulição, constantes críticas, entre outras e, empregando redes neurais, associar à formação ou não de azeótropos e às condições azeotrópicas. A estrutura básica do modelo consiste em empregar duas redes neurais (RN) diferentes. A primeira (RN1) correlaciona as propriedades dos componentes puros de modo a predizer se uma dada mistura binária é capaz ou não de formar um sistema azeotrópico. Se um dado sistema exibe comportamento azeotrópico, então, uma segunda rede (RN2) correlaciona as propriedades dos componentes puros dos sistemas binários com as condições do ponto azeotrópico (tipo, temperatura e composição). A estrutura do modelo é mostrada esquematicamente na Fig. 5.2.

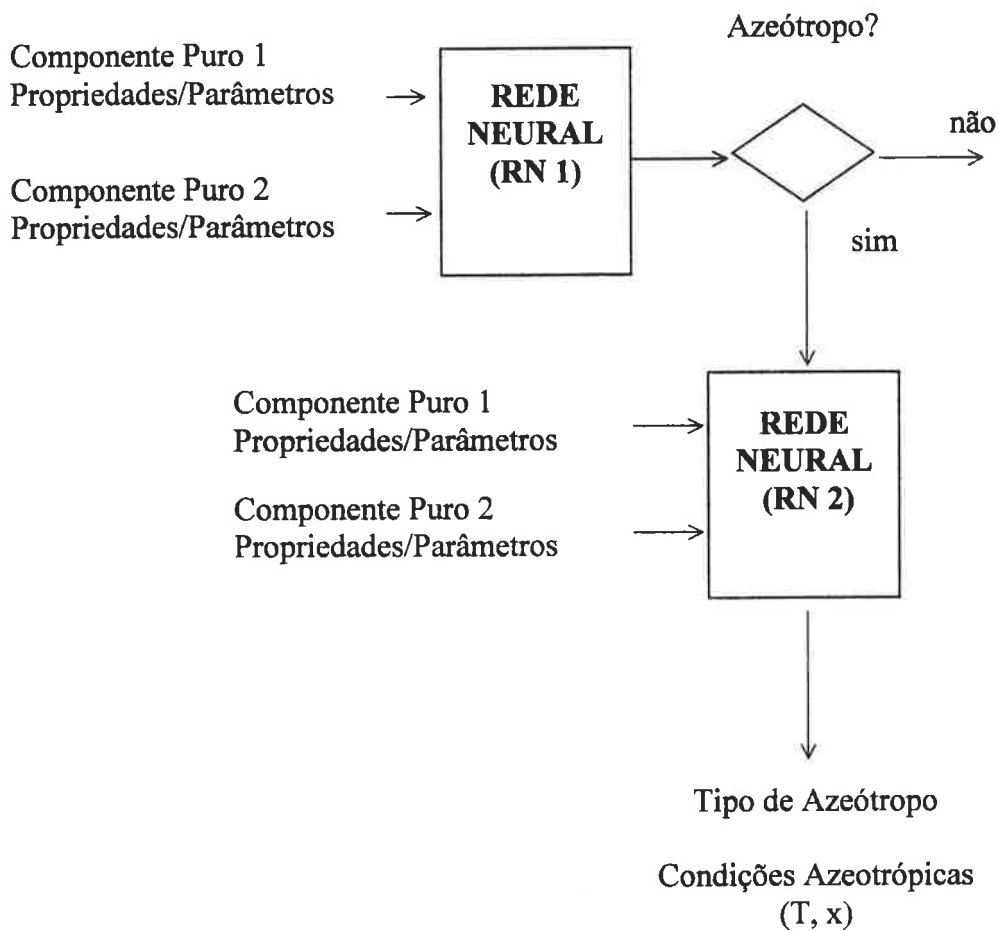


Fig. 5.2. Representação da Rede Neural - Predição de Azeótropo

A metodologia empregada consiste das seguintes etapas:

- a) Seleção de um conjunto adequado de propriedades de substâncias puras de modo a permitir uma boa caracterização do binário sob o ponto de vista do azeótropo;
- b) Elaboração de banco de dados com informações dos sistemas binários azeotrópicos (azeótropo, não azeótropo, tipo do azeótropo, temperatura, pressão e composição);
- c) Predição da formação de azeótropos:

- Seleção e treinamento da rede neural (RN 1);
- Verificação do modelo com dados novos;
- Análise dos resultados.

d) Predição das condições do ponto azeotrópico:

- Seleção e treinamento da rede neural (RN 2);
- Verificação do modelo com dados novos;
- Análise dos resultados.

Cada uma destas etapas será discutida em detalhes a seguir:

5.3.1. Seleção das variáveis da rede

Com base na teoria previamente descrita, foram selecionadas, para estudo, as seguintes propriedades para caracterização das substâncias puras e conseqüentemente associá-las a formação ou não de azeótropos e posterior predição do ponto azotrópico: Massa molecular (MM), temperatura de ebulição (T_{eb}), propriedades críticas (T_c , P_c e Z_c), fator acêntrico (ω), parâmetro de solubilidade (δ) e os descritores moleculares ou parâmetros solvatocrômicos de Abraham $\Sigma\alpha_2$, $\Sigma\beta_2$ e π .

Os parâmetros $\Sigma\alpha_2$ e $\Sigma\beta_2$ indicam a capacidade doadora ou receptora na formação de pontes de hidrogênio, de acordo com seu caráter ácido/base. Se ácido, $\Sigma\alpha_2 = 0$ - não doador; se base, $\Sigma\beta_2 = 0$ - não receptor.

O parâmetro π é um descritor dos efeitos polaridade e polarizabilidade combinados. São consideradas as interações dipolares do tipo dipolo - dipolo e dipolo - dipolo induzido.

O fator acêntrico foi introduzido de modo a considerar o grau de esfericidade da molécula, uma vez que as moléculas dos componentes de uma mistura azeotrópica geralmente são mais ou menos polares e não esféricas (Yashimoto & Mashiko, 1956).

As propriedades macroscópicas básicas foram obtidas em Reid et al. (1986), os parâmetros de solubilidade no banco de dados eletrônico DIPPR e os parâmetros solvatocrômicos de compilações da literatura (Abraham, 1993; Abraham et al., 1994).

A avaliação destes parâmetros em relação a sua importância para a predição azeotrópica será descrita na seção de seleção e treinamento da rede neural.

Além disso, para redução de dimensionalidade de um conjunto de dados multivariados, foi realizada uma avaliação das variáveis da rede neural utilizando a técnica de Análise de Componentes Principais (PCA). Vários métodos para selecionar um subconjunto de variáveis usando componentes principais são encontrados na literatura (Khattree & Naik, 2000). Para este trabalho, utilizou-se o programa SAS conforme descrito em Khattree & Naik, 2000. Os resultados podem ser vistos no ANEXO E.

Observa-se, a partir da tabela de autovalores da matriz correlação, que 10 componentes principais dos 20 totais seriam necessários para levar em conta mais que 95% da variação total dos dados. No entanto, avaliando a matriz correlação, observa-se que a maioria das correlações é muito fraca (a exceção das variáveis T_{eb1} e T_{cl}) e desde que baixa correlação entre as variáveis implica a ortogonalidade das variáveis no espaço original, fica difícil reduzir o espaço a uma dimensão menor. Assim, a matriz correlação indica se a redução de dimensão pode ser atingida. Resumindo, executar a análise de componentes principais permite, então conhecer qual o grau de redução possível e quais as variáveis que podem ser eliminadas, e, neste problema, corroborou com os resultados obtidos através do estudo original.

Assim, os resultados aqui apresentados, são aqueles obtidos através do estudo original.

5.3.2. Banco de dados de sistemas binários azeotrópicos

Os dados experimentais incluídos no banco de dados azeotrópicos foram obtidos a partir de compilações publicadas por Horsley (1973) e Gmehling et al. (1994) e de dados completos de equilíbrio líquido – vapor encontrados no banco de dados eletrônico TRC.

Foram observadas, na literatura, informações conflitantes em relação à existência ou não de azeótropo para um mesmo sistema. Assim, atenção especial foi dada à verificação e consistência da informação armazenada no banco de dados. Nos casos conflitantes, quando não foi possível a verificação diretamente a partir de curvas de equilíbrio líquido – vapor, foram considerados, como informação mais confiável, os dados referentes aos estudos mais recentes, levando em conta os avanços tecnológicos na aquisição de dados experimentais. Exemplos de tais sistemas podem ser vistos na Tabela 5.2:

Depois de exaustiva verificação da informação, o banco de dados final consta de cerca de 500 sistemas binários, os quais foram divididos em três grupos: um conjunto de treinamento (TREINO), um conjunto de teste (TESTE) e um conjunto de validação da rede neural (VALID), como visto adiante. TREINO foi usado para treinar a rede neural, isto é, obter os pesos da rede neural. TESTE foi usado para identificar o melhor conjunto de pesos, enquanto o terceiro conjunto, VALID, foi empregado como um conjunto completamente independente para validar a rede neural.

Tabela 5.2. - Comportamento Azeotrópico Discrepante - Exemplo

Sistema	Comportamento Azeotrópico	Referência	Fonte
i-Butanol (1) – m- Xileno (2)	Az	Garber, Bovkun and Etimova, Zh. Prikl. Khim. 35, 416 (1962); C.A. 56, 14997	Horsley/ Gmehling
	Az	Lecat, Ann. Chim., (12) 2, 158-202 (1947)	Horsley/ Gmehling
	Naz	Ryland, Am. Chem. J. 22, 390 (1899); Chem. News (London) 81, 15, 42, 50 (1908)	Horsley
Metil Acetato (1) – Ciclohexano (2)	Az	Hahn, Brenstoff-Chem. 35, 105 (1954)	Horsley
	Az	Nagata, J. Chem. Eng. Data 7, 461 (1962)	Horsley
	Naz	Lecat M., Ann. Soc. Sci. Bruxelles, 55B, 253 (1935)	Horsley
	Az	Tanaka T., Fluid Phase Equilibria 24, 187 (1985)	Gmehling
	Az	Nagata I., J. Chem. Eng. Data 7, 461 (1962)	Gmehling

5.3.3. Predição da formação de azeótropos

5.3.3.1. Seleção e treinamento da rede neural

Foi empregado um modelo de rede neural “feed-forward” constituído de três camadas e o algoritmo “back-propagation” para treinamento da rede.

Neste trabalho, as variáveis de entrada da rede neural foram as propriedades e parâmetros moleculares das substâncias puras e a variável de saída a informação azeotrópica (RN 1). Estas variáveis de entrada e saída constituem um conjunto de dados. Assim, cada conjunto de dados é um vetor contendo informação sobre as duas substâncias puras e sua capacidade para formar (ou não) uma mistura azeotrópica binária. A informação é disposta da seguinte forma: primeiro, são introduzidos os parâmetros e propriedades do componente mais volátil, seguidos daqueles dos componentes menos volátil (na mesma ordem) e, finalmente, o valor 1 para sistemas azeotrópicos ou 0 para sistemas não azeotrópicos.

A rede neural (RN 1) básica é mostrada esquematicamente na Figura 5.3 a seguir

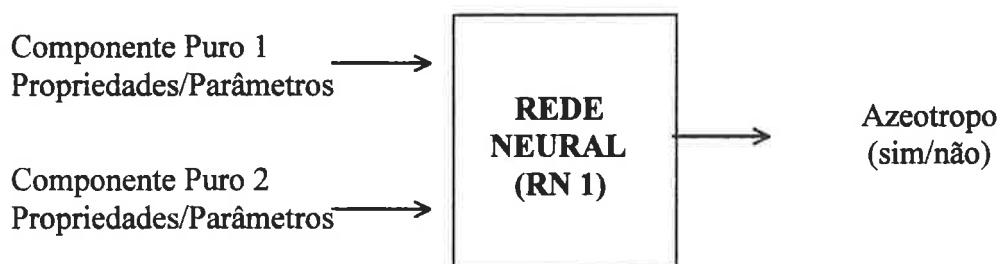


Figura 5.3. Representação Esquemática da Rede Neural - Predição de azeótropos

No desenvolvimento do modelo final da rede neural, vários casos com conjuntos de variáveis de entrada foram testados, empregando configurações com diferentes números de neurônio na camada oculta. Os dados foram designados para os conjuntos TREINO e TESTE de diferentes maneiras e o efeito da variação do número de apresentações durante o treinamento foi examinado. A melhor configuração de rede neural, em cada caso, foi a que forneceu o menor erro nas predições para o TESTE.

Em todos os casos avaliados, o acerto entre os valores experimentais e calculados para todas as variáveis no conjunto TREINO foi de 100%, como esperado, uma vez que normalmente o conjunto TREINO é sempre ajustado. Então, na fase de treinamento, o mais importante é a predição para conjunto TESTE.

Os resultados dos casos estudados, são apresentados a seguir, para a melhor configuração de rede obtida para conjunto de propriedades e parâmetros das substâncias puras que representam a formação azeotrópica.

Caso I:

Variáveis de entrada: MM, Teb, Tc, Pc, Zc, α , β e π
(para cada componente)

Neurônios na camada interna: 12

Pontos no conjunto de aprendizagem: 215

Pontos no conjunto de teste: 66

Os resultados obtidos mostram um bom ajuste dos pontos de treinamento com erro de $1,73 \times 10^{-1}$, mas o ajuste do teste apresenta um erro muito grande, 5,19, além de oscilar muito, sempre em torno de um mesmo conjunto de valores. A discriminação do conjunto de pontos do teste apresenta um acerto da ordem de apenas 76%, bem inferior ao obtido para um conjunto menor de dados. Estes resultados são mostrados na Figura 5.4 a seguir.

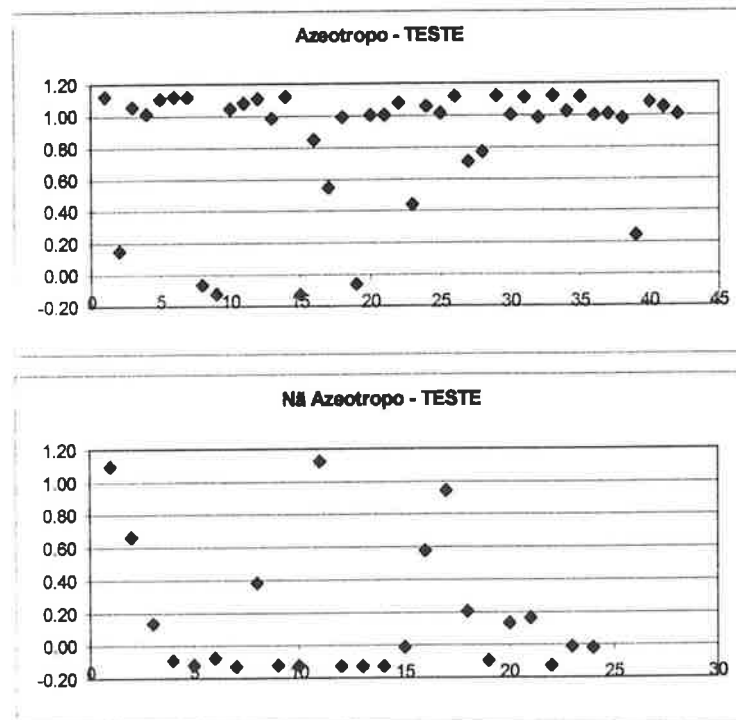


Fig. 5.4. Resultado da Rede Neural – Predição de Azeótopos - Caso I

Neste ponto, após vários testes adicionais, creditou-se a grande diferença observada a um problema de domínio das variáveis, isto é todas as variáveis do conjunto TESTE podem não estar completamente inseridas no domínio das

variáveis do conjunto TREINO da rede. O principal problema é a dificuldade em averiguar exatamente quais conjuntos de dados do TESTE estão fora do domínio do TREINO, uma vez que as variáveis de entrada envolvidas são muitas (16), o que impossibilita uma rápida visualização ou análise.

Também foi verificada uma certa inconsistência em relação a sistemas de séries de isômeros, quando nem todos se comportam da mesma forma frente à formação de azeótropos e fazem parte de conjuntos de dados diferentes, isto é, alguns estão no TREINO e outros no TESTE. Como as propriedades dos isômeros são muito parecidas, a discriminação fica prejudicada. A solução para este problema foi tentar encontrar alguma informação (propriedade/parâmetro) a partir da qual os isômeros pudessem ser facilmente identificados ou diferenciados. Foi sugerida a utilização de índices de conectividade, que fornecem indicações sobre a forma e tamanho das moléculas e/ou índices de valência, os quais dão uma medida da natureza eletrostática da molécula. Estes índices foram calculados para todos os isômeros da lista de substâncias estudadas através do programa MOLECULAR MODELING PRO Computational Chemistry Program da Window Chem Software Inc. Como resultado, não se obteve qualquer informação relevante que pudesse ser utilizada na estrutura da rede neural como agente diferenciador dos isômeros.

Assim, antes de passar para a solução do problema, decidiu-se trabalhar com o conjunto completo de dados azeotrópicos e não azeotrópicos disponíveis na literatura, para o grupo de substâncias selecionadas para este estudo, totalizando 490 dados com a informação desejada. A distribuição destes pontos entre os conjuntos de TREINO, TESTE e VALID não mais foi feita através da utilização do programa PREP2.FOR, como anteriormente, e sim de acordo com a análise criteriosa do domínio de cada variável nos conjuntos de dados.

A metodologia empregada foi a de analisar grupos de conjuntos de dados de acordo com os componentes constituintes do binário, por exemplo considerando o n-hexano como componente 1, selecionou-se todos os binários nos quais o n-

hexano aparece como tal e verificou-se o intervalo de domínio de todas variáveis do componente 2 simultaneamente em todos os sistemas em que figuram com tal. Este procedimento foi executado para todos constituintes dos sistemas binários apresentados e checados fazendo-se o procedimento inverso, isto é, fixando o componente 2 e verificando o domínio das variáveis do componente 1. Desta forma, toda vez que estes critérios eram satisfeitos para um certo conjunto de dados, o mesmo era retirado do conjunto total e adicionado ao conjunto de TESTE/VALID. Embora seja um procedimento demorado e exaustivo, pode-se dizer, com alguma certeza que todos os sistemas selecionados para os subconjuntos TESTE/VALID encontram-se bem representados pelos sistemas constituintes do conjunto TREINO.

Finalmente, os 490 sistemas foram divididos entre 405 dados para o conjunto de TREINO, 43 para o conjunto de TESTE e 42 para o conjunto de VALID com base na análise de clusters descritos acima para assegurar distribuição sobre o domínio dentro de cada subconjunto. Esta estratégia, essencial para a validade do modelo da rede, atribuiu todos os pares nos quais um dos componentes era água (43 sistemas) ou clorofórmio (27 sistemas), todos os sistemas azeotrópicos heterogêneos (33 sistemas, todos de máxima pressão) a todos os sistemas azeótropos de mínima pressão (4 sistemas, todos homogêneos) ao conjunto de treinamento.

Ainda com o objetivo de verificar a validade da informação, todos os dados referentes à formação de azeótropos foram, uma vez mais, checados nas diversas fontes utilizadas para sanar qualquer dúvida e/ou algum erro de transcrição por ventura existente.

Com estes novos conjuntos de dados, mais uma vez foi verificada a estrutura da rede, analisando-se todas as possíveis alternativas em relação às propriedades das substâncias puras usadas com variáveis de entrada da rede, tomando-se como ponto de partida, os resultados obtidos no caso I mostrado acima. Assim:

Caso II:

Variáveis de entrada: MM, T_{eb} , T_c , P_c , Z_c , α , β e π
(para cada componente) – total: 16

Neurônios na camada interna: 11

Pontos no conjunto de aprendizagem: 405

Pontos no conjunto de teste: 43

A concordância entre os valores experimentais e calculados foi de 90,7% com 4 pontos não discriminados. Os resultados são mostrados na Fig. 5.5.

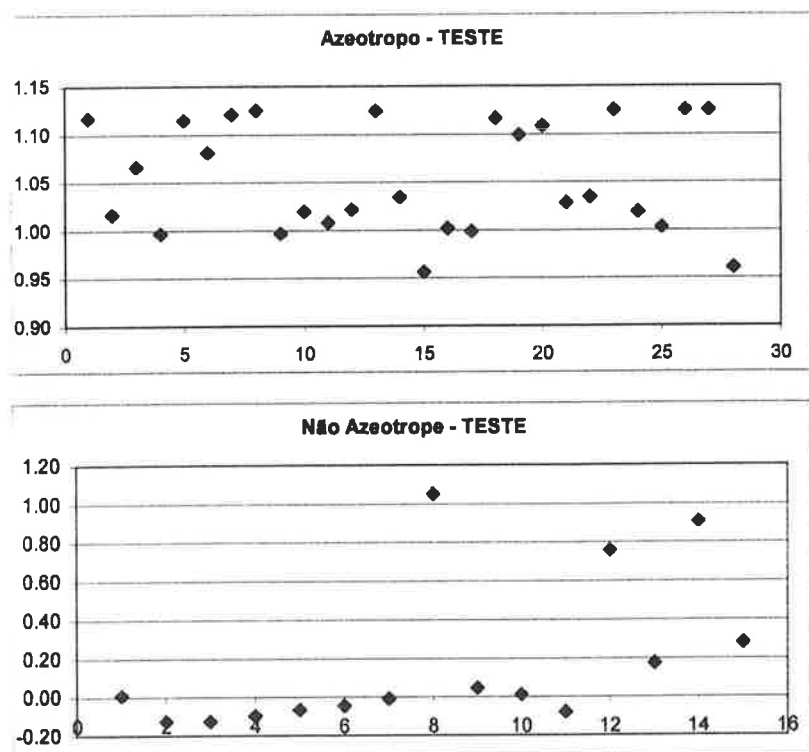


Fig. 5.5.- Resultado da Rede Neural – Predição de Azeótropos – Caso II

Caso III:

Variáveis de entrada: MM, T_{eb} , T_c , P_c , Z_c , α , β , π e δ
(para cada componente) – total: 18

Neurônios na camada interna: 13

Pontos no conjunto de aprendizagem: 405

Pontos no conjunto de teste: 43

A concordância entre os valores experimentais e calculados foi de 81,4% com 8 pontos não discriminados. Se for considerado o mesmo número de neurônios na camada oculta do caso anterior ($NH = 11$), o acerto é da ordem de 90,4%, no entanto são duas variáveis a mais em relação ao mesmo caso anterior. Os resultados são mostrados na Fig. 5.6.

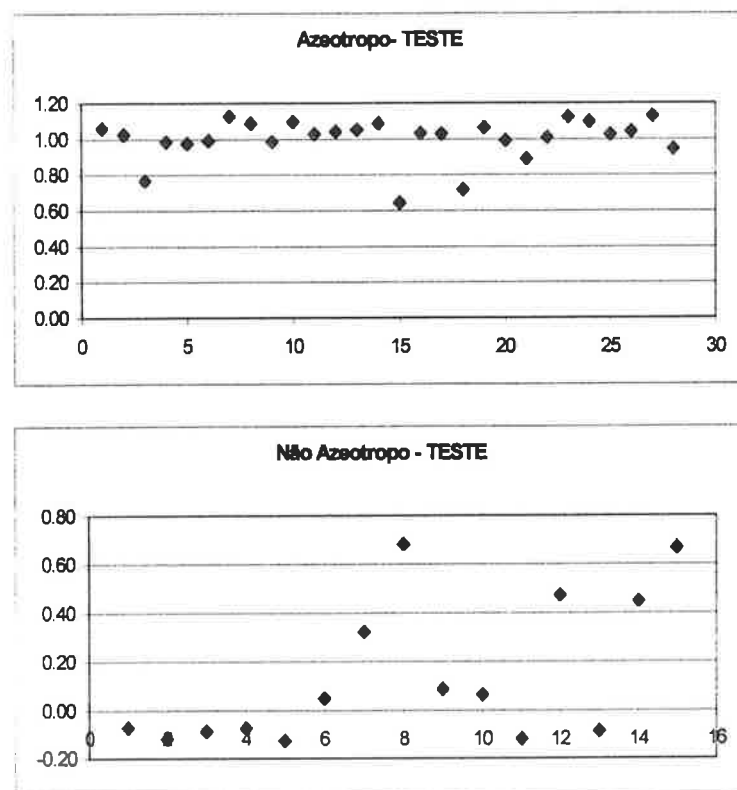


Fig. 5.6. - Resultado da Rede Neural – Predição de Azeótropos – Caso III

De acordo com os resultados obtidos nestes casos, a melhor escolha seria o correspondente ao Caso II, uma vez que apresenta menor número de variáveis envolvidas e, de acordo com o critério adotado (menor erro para o TESTE), menor número de neurônios na camada oculta.

Apesar dos bons resultados obtidos até aqui, verificou-se a influência da introdução do fator acêntrico e do parâmetro de solubilidade na predição azeotrópica. Assim:

Caso IV:

Variáveis de entrada: MM, Teb, Tc, Pc, Zc, ω , α , β e π
(para cada componente)- total: 18

Neurônios na camada interna: 11

Pontos no conjunto de aprendizagem: 405

Pontos no conjunto de teste: 43

A concordância entre os valores experimentais e calculados foi de 93,02% com 3 pontos não discriminados. Pode-se observar que o grau de acerto neste caso é maior que o do caso II, no entanto o número de variáveis é maior. Os resultados são mostrados na Fig. 5.7.

Caso V:

Variáveis de entrada: MM, Teb, Tc, Pc, Zc, ω , α , β , π e δ
(para cada componente) – total: 20

Neurônios na camada interna: 12

Pontos no conjunto de aprendizagem: 405

Pontos no conjunto de teste: 43

A concordância entre os valores experimentais e calculados foi de 90,7% com 3 pontos não discriminados. Os resultados são mostrados na Fig. 5.8.

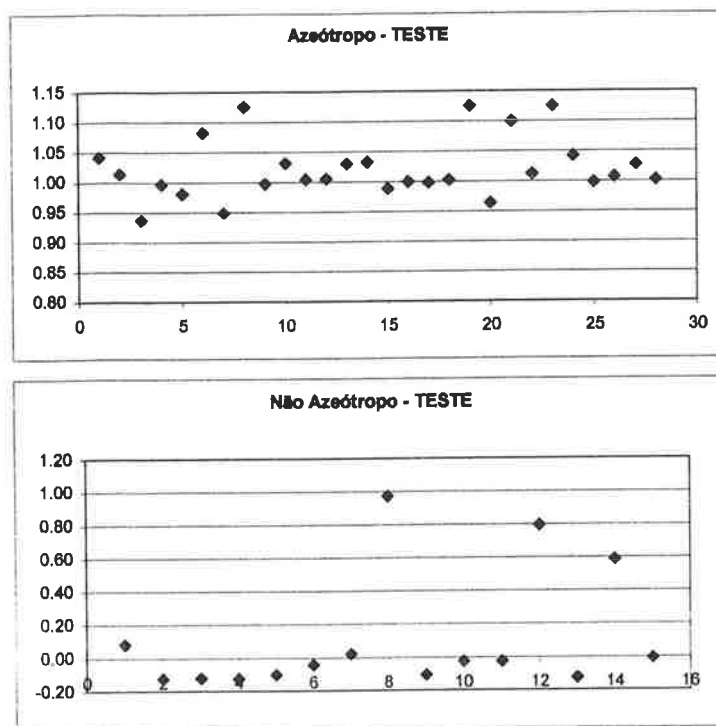


Fig. 5.7. - Resultado da Rede Neural – Predição de Azeótropos – Caso IV

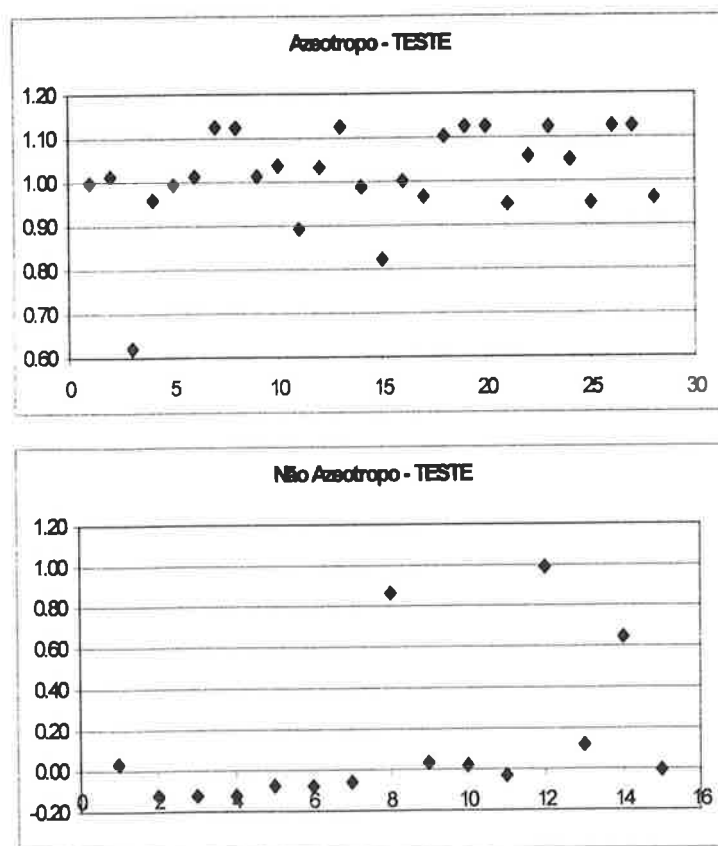


Fig. 5.8. - Resultado da Rede Neural – Predição de Azeótropos – Caso V

De acordo com os resultados obtidos e considerando o menor número de variáveis e menor número de neurônios na camada oculta, a melhor escolha é o Caso II, isto é, um conjunto de oito variáveis de entrada para cada componente puro da mistura binária: MM , T_{eb} , T_c , P_c , z_c , π_2 , $\Sigma\alpha_2$, e $\Sigma\beta_2$. Embora T_c seja sabidamente correlacionada empiricamente com MM e T_{eb} (Reid et al., 1988), a correlação tem um erro relativamente grande (pelo menos, para os objetivos deste trabalho). De fato, a exclusão de MM ou T_{eb} do conjunto das variáveis de entrada resultou em redes neurais com desempenho significativamente mais pobre.

Como notado acima, o sucesso na obtenção de uma rede neural robusta depende fortemente da escolha das variáveis envolvidas, bem como da qualidade e domínio dos dados usados para efeito de treinamento (Chan & Nascimento, 1994).

Para todos os casos estudados, sobreajuste foi evitado selecionando a rede neural que apresentou o menor número possível de interações e o menor número de neurônios. O critério principal para seleção da rede mais adequada foi o menor erro para o conjunto de teste (TESTE). Além disso, valores da variável de saída maiores ou iguais a 0,8 foram usados para discriminar um azeótropo e menores ou iguais a 0,2 para sistemas não azeotrópicos. Valores entre estes dois limites foram classificados como casos para os quais, a predição não foi possível empregando o modelo de redes neurais proposto.

A Tabela 5.3 mostra os resultados para a predição azeotrópica para o conjunto TESTE, considerando a configuração do caso II.

Tabela 5.3. - Comparação entre valores experimentais e preditos utilizando modelo de redes neurais e método UNIFAC – TESTE

Componente 1	Componente 2	Valor Experimental	RN	UNIFAC
Metil Acetato	Metil-Ciclopentano	Az	Az	Az
Etanol	Tolueno	Az	Az	Az
Hexano	Butanol	Az	Az	Az
Etanol	p-Xileno	Naz	Naz	Naz
Etanol	Heptano	Az	Az	Az
Acetona	n-Propanol	Naz	Naz	Naz
Acetona	Isobutanol	Naz	Naz	Naz
Acetona	Hexano	Az	Az	Az
2-Butanona	Álcool Alílico	Naz	Naz	Naz
Álcool Alílico	Clorobenzeno	Az	Az	Az
Metil-Ciclopentano	Isopropanol	Az	Az	Az
Isopropanol	Etilbenzeno	Naz	Naz	Naz
1,3-Ciclohexadieno	n-Propanol	Az	Az	Az
n-Propanol	Metil-Ciclohexano	Az	Az	Az
Etil Acetato	Fluorbenzeno	Naz	Naz	Az
Butanol	p-Xileno	Az	Az	Az
Séc-Butanol	Tolueno	Az	Az	Az
Isobutanol	Etilbenzeno	Az	Az	Az
3-Pentanona	3-Pentanol	Naz	Naz	Naz
Ciclohexano	2-Pentanol	Naz	Az	Az
Metil-Ciclohexano	2-Pentanol	Az	Az	Az
Benzeno	3-Pentanol	Naz	Naz	Az
Heptano	3-Pentanol	Az	Az	Az
Metil-Ciclohexano	Tolueno	Naz	Naz	Az
n-Propanol	Clorobenzeno	Az	Az	Az
n-Propanol	Tolueno	Az	Az	Az
2-Butanona	Tolueno	Naz	Naz	Naz
Álcool Alílico	Tolueno	Az	Az	Az
Metanol	Clorobenzeno	Naz	NP	Naz
Metil Acetato	Etanol	Naz	Naz	Naz
Metil-Ciclopentano	sec-Butanol	Az	Az	Az
Pentano	n-Propanol	Naz	Az	Az
Tolueno	2-Pentanol	Az	Az	Az
Ciclohexano	sec-Butanol	Az	Az	Az
Benzeno	2-Pentanona	Naz	NP	Naz
Etil Acetato	Benzeno	Az	Az	Az
Álcool Alílico	3-Pentanona	Az	Az	Az
1,3-Ciclohexadieno	Álcool Alílico	Az	Az	Az
Butanol	m-Xileno	Az	Az	Az
Etanol	Fluorbenzeno	Az	Az	Az
Ciclohexano	Isobutanol	Az	Az	Az
Ciclohexano	n-Propanol	Az	Az	Az
Isobutanol	m-Xileno	Az	Az	Az

Notas: Naz – não azeótropo; Az – Azeótropo; NP – não predito

5.3.3.2. Validação da Rede Neural e Análise dos Resultados

Após obtenção da rede neural treinada e do conjunto de pesos apropriados que minimizam o erro quadrado médio, avaliações foram feitas de modo a checar o desempenho da rede. Para isto foi usado o terceiro grupo de dados, aqui chamado de conjunto de validação (VALID). Os resultados foram muito satisfatórios com 88% de acerto na predição de formação azeotrópica. A Tabela 5.3 mostra a comparação entre os valores experimentais e preditos para VALID utilizando a rede neural apresentada e o método UNIFAC. Examinando a Tabela 3, observa-se que, dos 42 sistemas binários em VALID, apenas dois não foram preditos corretamente pelo modelo de redes neurais (Etil Acetato – 2-Butanona e Acetona – Benzeno), enquanto que para os sistemas n-Hexano – 2-Pentanona, Metanol – 2-Pentanona e Etanol – 3-Pentanona, não foi possível uma boa discriminação, de acordo com os critérios adotados, e para estes sistemas pode-se dizer a rede neural não foi capaz de predição.

Pode ser observado que, nos três casos para os quais a rede neural falhou na predição, são sistemas onde um dos constituintes é 2- ou 3-Pentanona. Uma explicação razoável pode ser dada analisando o conjunto TREINO, onde se verificam algumas particularidades em relação a sistemas formados por estes compostos e, por exemplo, uma família de álcoois, como mostrado na Tabela 5.5.

De acordo com Swietoslowski (1959), uma série de azeótropos é formada por um componente comum e membros sucessivos de uma série de homólogos, isômeros ou substâncias relacionadas. Em geral, uma vez que propriedades físico-químicas de homólogos de uma dada série variam de maneira ordenada, seria esperado que o comportamento azeotrópico de tal série de azeotrópico exibissem a mesma regularidade. Desta forma, seria esperado que para os sistemas binários da Tabela 5.5 apresentassem as mesmas características, mas isto não ocorre, como pode ser visto para as misturas binárias das pentanonas com álcoois alifáticos quando comparadas com as misturas dos mesmos álcoois com 2-Butanona.

Tabela 5.4. - Comparação entre valores experimentais e preditos utilizando modelo de redes neurais e método UNIFAC – VALID

Componente 1	Componente 2	Valor Experimental	NN	UNIFAC
n-Pentano	Etanol	Az	Az	Az
Ciclohexano	Butanol	Az	Az	Az
Etanol	Clorobenzeno	Naz	Naz	Az
Etil Acetato	2-Butanone	Az	Naz	Az
2-Butanona	Metil Ciclohexano	Az	Az	Az
Etanol	n-Propanol	Naz	Naz	Naz
Metanol	m-Xileno	Az	Az	Naz
Acetona	Butanol	Naz	Naz	Naz
Acetona	Benzeno	Naz	Az	Naz
Acetona	1,3-Ciclohexadieno	Az	Az	Az
Etil Acetato	Álcool Alílico	Naz	Naz	Naz
Álcool Alílico	n-Heptano	Az	Az	Az
n-Propanol	sec-Butanol	Naz	Naz	Naz
n-Propanol	Isobutanol	Naz	Naz	Naz
n-Propanol	p-Xileno	Az	Az	Az
2-Butanone	Fluorbenzeno	Az	Az	Az
Metil Ciclopentano	Isobutanol	Az	Az	Az
Isopreno	Etanol	Az	Az	Az
Isopreno	Metanol	Az	Az	Az
n-Hexano	2-Pentanona	Naz	NP	Az
Etanol	Butanol	Naz	Naz	Naz
Isopropanol	Álcool Alílico	Naz	Naz	Naz
Ciclopentano	Metil Acetato	Az	Az	Az
Metil Acetato	Ciclohexano	Az	Az	Az
2-Butanona	sec-Butanol	Naz	Naz	Naz
n-Pentano	Benzeno	Naz	Naz	Naz
n-Pentano	Ciclohexano	Naz	Naz	Naz
Benzeno	Metil Ciclohexano	Naz	Naz	Naz
n-Hexano	Tolueno	Naz	Naz	Naz
n-Heptano	Etilbenzeno	Naz	Naz	Naz
Tolueno	p-Xileno	Naz	Naz	Naz
Metanol	Isobutanol	Naz	Naz	Naz
Metanol	2-Pentanona	Naz	NP	Naz
Etanol	3-Pentanona	Az	NP	Az
Álcool Alílico	Etilbenzeno	Az	Az	Naz
Acetonitrila	m-Xileno	Naz	Naz	Naz
Metil Acetato	n-Propanol	Naz	Naz	Naz
Methyl Acetate	Toluene	Naz	Naz	Naz
Isopropyl Alcohol	Butyl Alcohol	Naz	Naz	Naz
2-Butanone	Chlorobenzene	Naz	Naz	Naz
Ethyl Acetate	3-Pentanone	Naz	Naz	Naz
Séc-Butyl Alcohol	Isobutyl Alcohol	Naz	Naz	Naz

Notas: Naz – não azeótropo; Az – Azeótropo; NP – não predito

Tabela 5.5. Dados azeotrópicos experimentais para sistemas formados por álcoois e cetonas

	2-Butanona	2-Pentanona	3-Pentanona
Metanol	Az	Naz	Naz
Etanol	Az	Az	Az
n-Propanol	Naz	Az	Az
Isopropanol	Az	Naz	Naz
Butanol	Naz	Naz	Naz
Isobutanol	Naz	Az	Az
sec-Butanol	Naz	-	Az
terc-Butanol	Az	-	Naz

Notas: Naz – não azeótropo

Az – Azeótropo

Pode-se concluir qualquer que seja o método baseado nestas idéias está sujeito a falhas no que diz respeito à predição azeotrópica para estes tipos de misturas binárias. Entretanto, o método UNIFAC faz a correta predição para estes casos.

5.3.4. Predição das condições do ponto azeotrópico

Nesta etapa do trabalho, dando continuidade a predição do azeótropo, uma vez determinado se um dado sistema forma ou não um azeótropo, é necessário determinar as condições azeotrópicas, isto é, predizer a temperatura e a composição na qual ocorre o azeótropo. A proposta tem a mesma base anterior, ou seja, a partir de propriedades das substâncias puras e utilizando redes neurais, determinar o tipo, a temperatura e a composição do azeótropo à pressão atmosférica. A Fig. 5.9 mostra esquematicamente o procedimento adotado.

Inicialmente, partiu-se do mesmo conjunto de propriedades e parâmetros dos componentes puros usados na primeira parte do trabalho. No entanto, para este caso, a introdução do fator acêntrico apresentou grandes melhorias no resultado da predição. Assim, as informações de entrada da RN2 são: propriedades macroscópicas: MW, T_{eb} , T_c , P_c , Z_c , w e os parâmetros moleculares de Abraham: α, β e π . A saída da RN 2 é o tipo de azeótropo: heterogêneo ou

homogêneo, máximo ou mínimo, codificado por números inteiros e as condições azeotrópicas (temperatura e composição à pressão atmosférica).

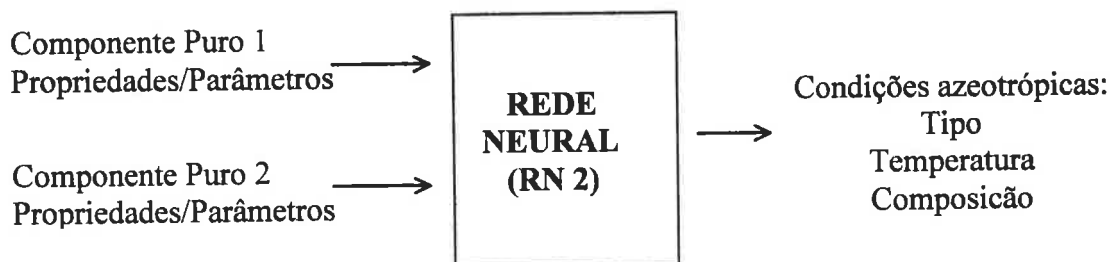
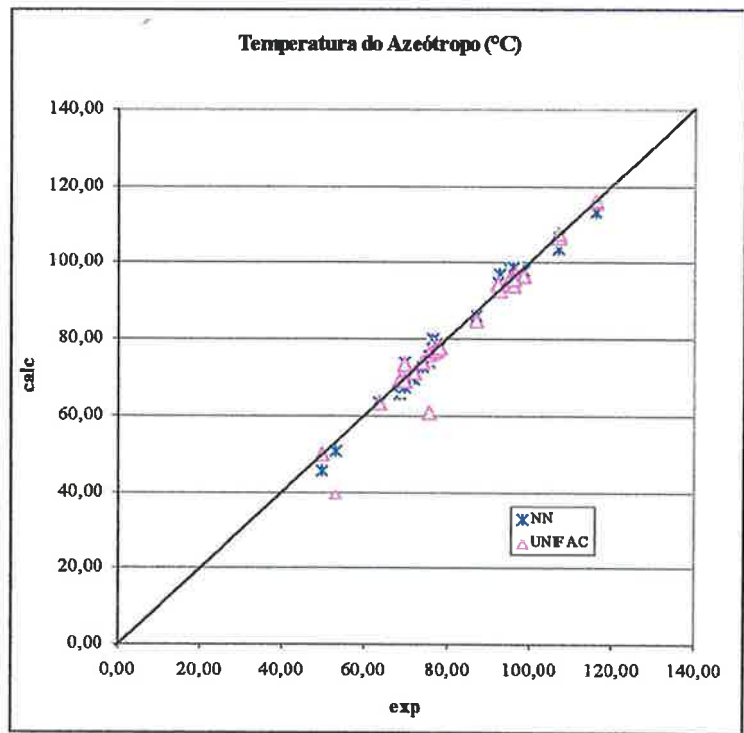
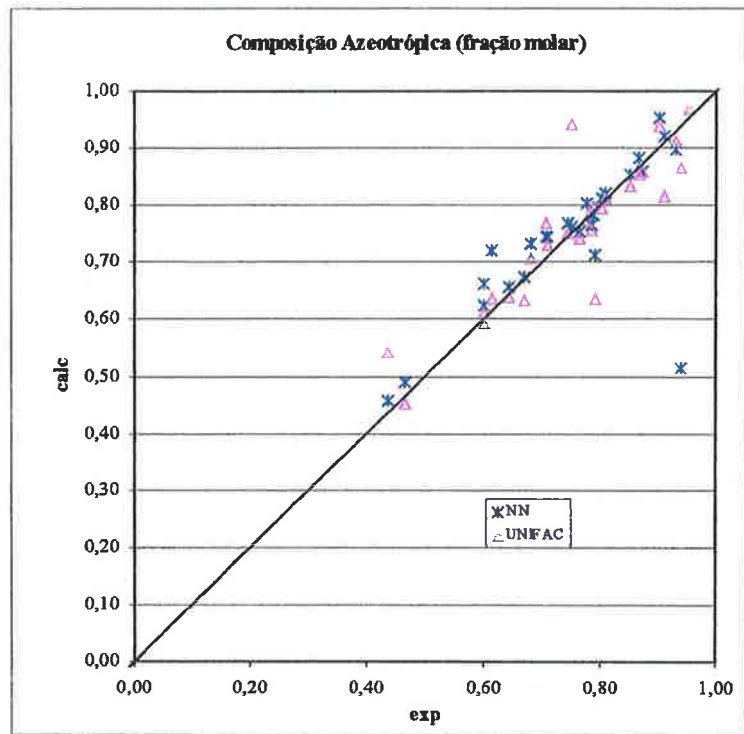


Fig. 5.9. - Representação Esquemática da Rede Neural (RN2)
Predição das condições azeotrópicas

O banco de dados nesta parte do trabalho é constituído dos sistemas discriminados como azeótropos na etapa anterior, sendo então divididos em 185 conjuntos de dados para a etapa de treinamento, 28 sistemas no conjunto de teste e 16 sistemas binários no conjunto de validação. Não foi encontrada qualquer dificuldade na convergência da rede neural durante o processo de treinamento. A discriminação do tipo de azeótropo, usando uma rede com 11 neurônios na camada interna, foi de 100%.

A concordância entre os valores experimentais e os valores preditos de temperatura e composição também foi muito boa. As Figuras 5.10.a e 5.10.b mostram os resultados. Pode ser observado que a temperatura azeotrópica foi muito bem representada e na predição da composição apenas um ponto não foi discriminado, enquanto os demais foram satisfatoriamente representados. Estas figuras também a comparação com os valores preditos utilizando o método UNIFAC. Em geral, os casos estudados foram melhores representados pelo modelo de redes neurais proposto. Estes resultados mostraram não ser necessário a introdução de parâmetros adicionais para as substâncias puras em relação à predição das condições azeotrópicas.



Figuras 5.10.a/b. Predição das condições azeotrópicas a 1 atm
 Comparação entre Redes Neurais e UNIFAC - Conjunto de Teste

5.4. – Conclusões

A predição da formação de azeótropos em misturas binárias usando apenas propriedades macroscópicas e microscópicas de substâncias puras, constituintes do sistema, sem qualquer consideração explícita da não-idealidade da mistura, é um desafio interessante para metodologia de redes neurais. Este trabalho demonstra que um modelo simples de redes neurais com variáveis de entradas fisicamente adequadas para representar os componentes puros pode fornecer uma boa correlação para a predição da ocorrência (ou ausência) de formação azeotrópica, bem como as condições azeotrópicas para um grande número de sistemas binários. Em um relativamente pequeno número de situações, o modelo falha ao fazer uma predição exata com relação à formação de azeótropos. Todos estes correspondem predominantemente àqueles casos nos quais moléculas estruturalmente homólogas exibem um comportamento azeotrópico distinto. Comparação entre o modelo de redes neurais e o modelo UNIFAC mostra que o modelo de redes neurais pode prever as condições azeotrópicas melhor que o método UNIFAC, porém resultados conclusivos necessitam novas avaliações com mais e diversificados sistemas binários.

6. CONCLUSÕES

A aplicação de redes neurais é uma ferramenta muito atrativa para lidar com problemas de engenharia química, principalmente no que diz respeito ao tratamento de problemas complexos, multivariáveis, não-lineares e aqueles que não dispõem de um bom modelo fenomenológico para bem representá-lo. Aliado a sua capacidade em ajustar dados, a grande vantagem em utilizar modelos de redes neurais é não ser necessário um conhecimento profundo das leis físicas e químicas que definem o sistema em estudo.

Este trabalho mostra com sucesso uma variedade de aplicações práticas e específicas de modelos de redes neurais a uma unidade industrial, tais como: análise de dados para detecção de “outliers”, modelagem, simulação e otimização de processos e predição de propriedades termodinâmicas, indispensáveis para o conhecimento do comportamento da operação da planta.

No caso estudado, foi possível simular o processo industrial, principalmente devido à indisponibilidade de dados termodinâmicos necessários para utilização de ferramentas convencionais. Os resultados da simulação do modelo global foram coerentes com o comportamento da planta, evidenciando que a metodologia desenvolvida pode ser empregada com segurança para a representação do processo de produção de isopreno.

A partir do modelo global gerado foram realizados estudos de análise de sensibilidade e otimização. Os resultados obtidos nos estudos fornecem sugestões de condições operacionais possíveis de aplicação na unidade, ressaltando o aspecto prático do estudo desenvolvido.

Aspectos relevantes das diversas aplicações de redes neurais foram consideradas nas conclusões específicas apresentadas nos capítulos 4 e 5.

7. SUGESTÕES DE CONTINUIDADE DO TRABALHO

Algumas sugestões para continuidade do trabalho são citadas a seguir:

- ✓ Sistematizar e automatizar os procedimentos de tratamento de dados para eliminação de erros grosseiros e sistemáticos.
- ✓ Automatizar o procedimento de estabelecimento de períodos de operação em regime estacionário.
- ✓ Empregar metodologia de modelagem híbrida, empregando modelos fenomenológicos para os subsistemas para os quais se dispõem de dados termodinâmicos confiáveis e modelos de redes neurais para os demais sistemas.
- ✓ Realizar estudos de análise de sensibilidade para as variáveis do processo não consideradas neste estudo.
- ✓ Estabelecer funções objetivas econômicas mais específicas para otimização do processo.
- ✓ Ampliar o banco de dados de substâncias puras, tanto em número quanto em diversidade de funções químicas, de modo a generalizar a metodologia proposta para predição azeotrópica.
- ✓ Empregar a metodologia de mapas de curvas residuais para avaliação da seção de recuperação de solvente da unidade em estudo, principalmente no que diz respeito ao teor de água no solvente circulante.

ANEXO A - SELEÇÃO DO NÚMERO DE NEURÔNIOS DAS REDES NEURAIS

A influência do número de neurônios na camada oculta (NH) e do número de iterações no erro residual, objetivando a seleção da melhor configuração das redes, é apresentada nas figuras a seguir, para cada uma das redes individuais representativas de cada subsistema.

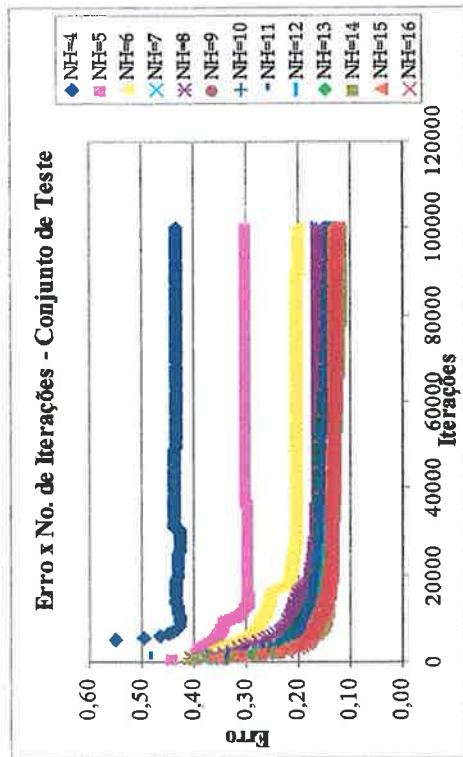
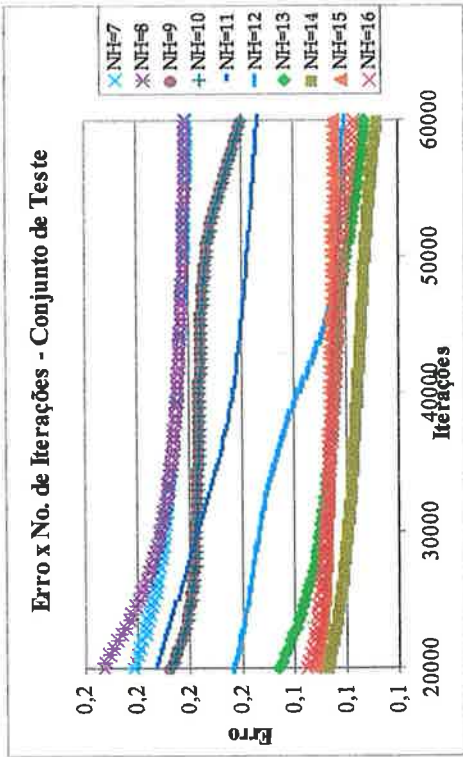


Fig. A.1 – Seleção do Número de Neurônios da REDE 1

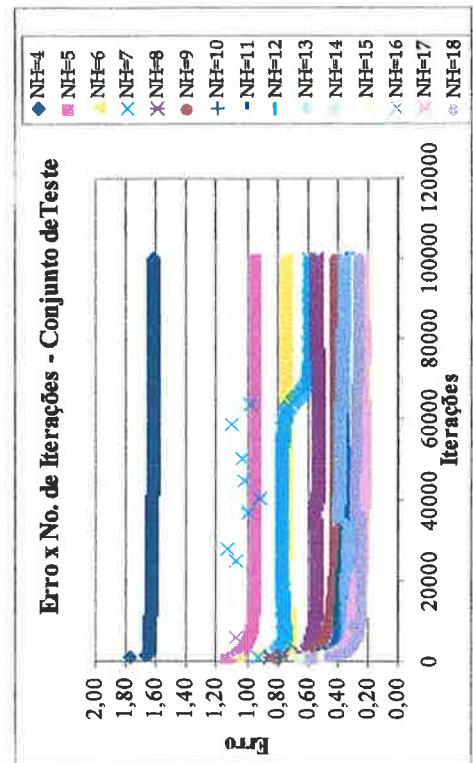
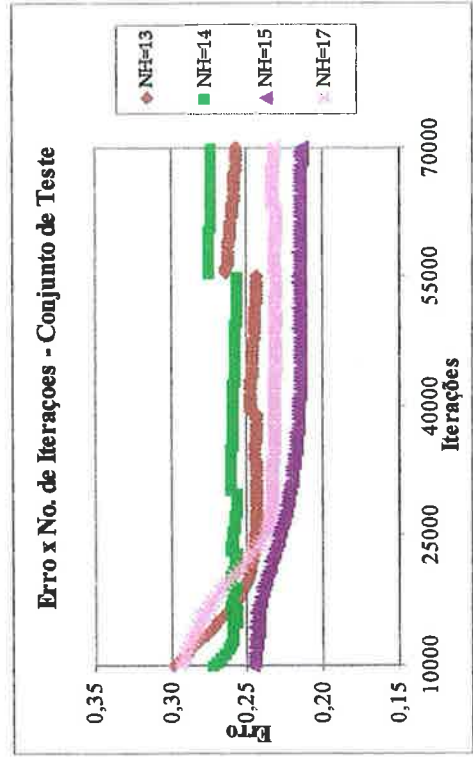


Fig. A.2. – Seleção do Número de Neurônios da REDE 2

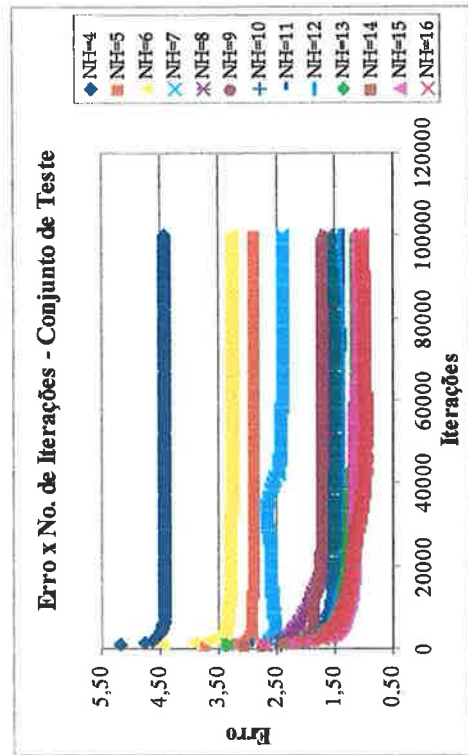
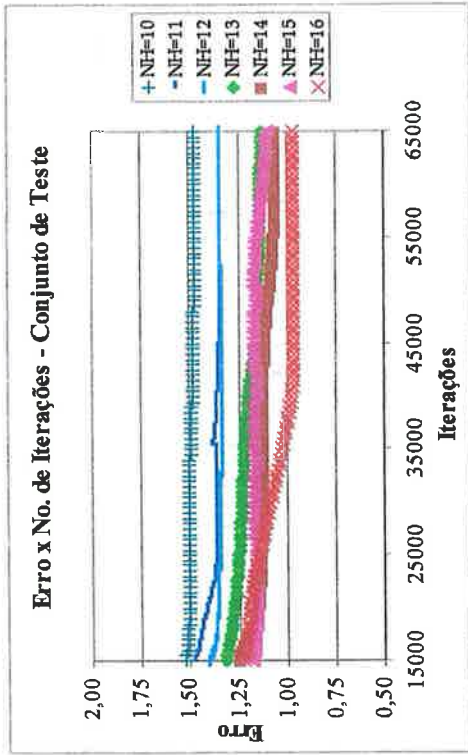


Fig. A.3. – Seleção do Número de Neurônios da REDE 3

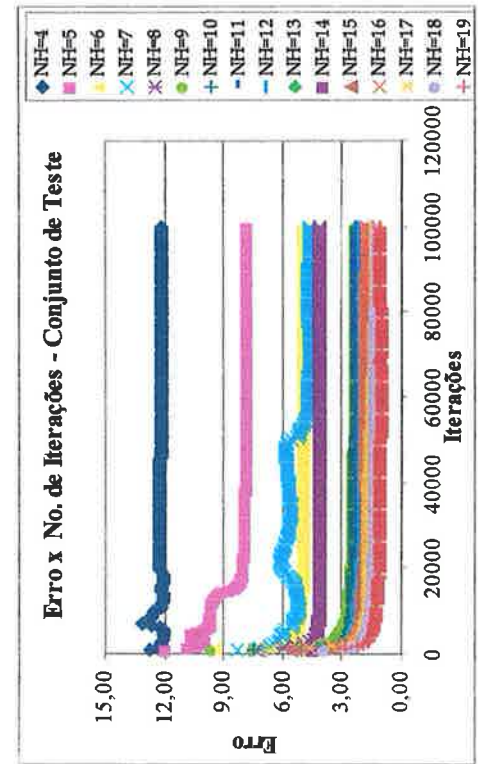
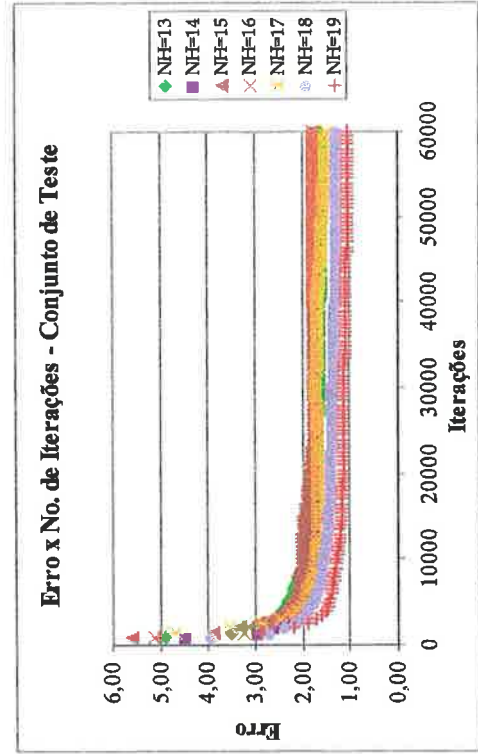


Fig. A.4. – Seleção do Número de Neurônios da REDE 4

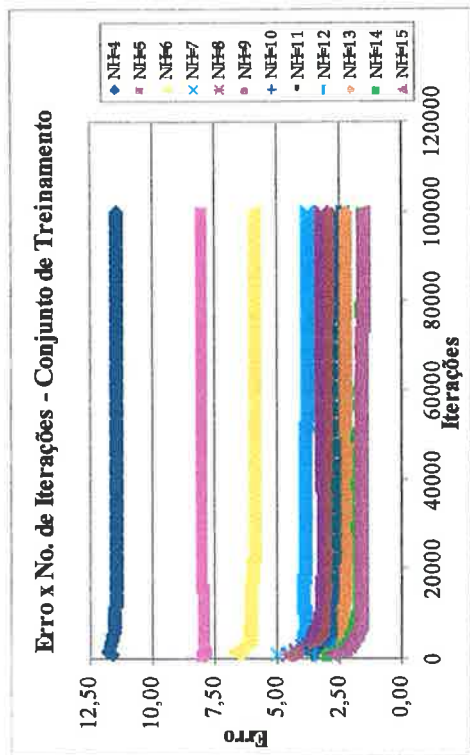
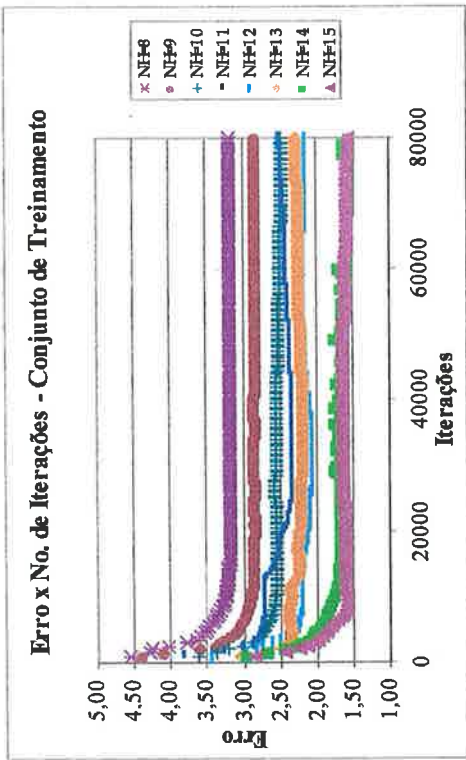


Fig. A.5. – Seleção do Número de Neurônios da REDE 5

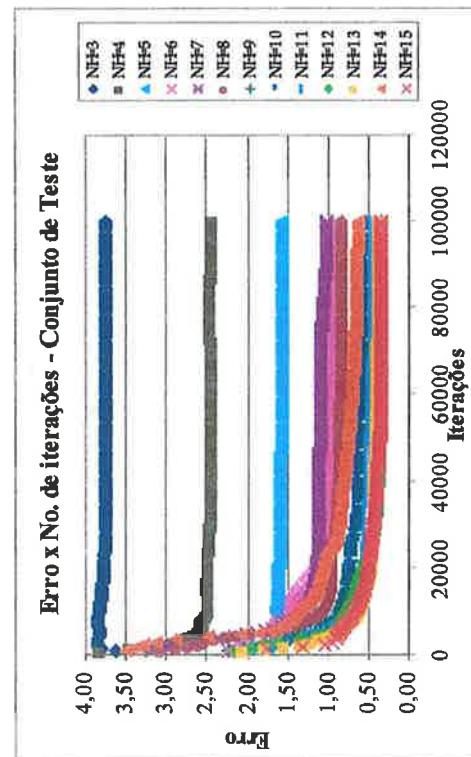
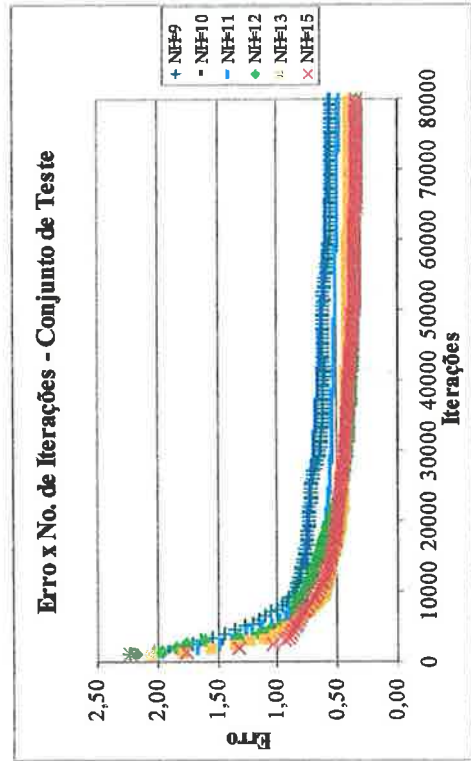


Fig. A.6. – Seleção do Número de Neurônios da REDE 6

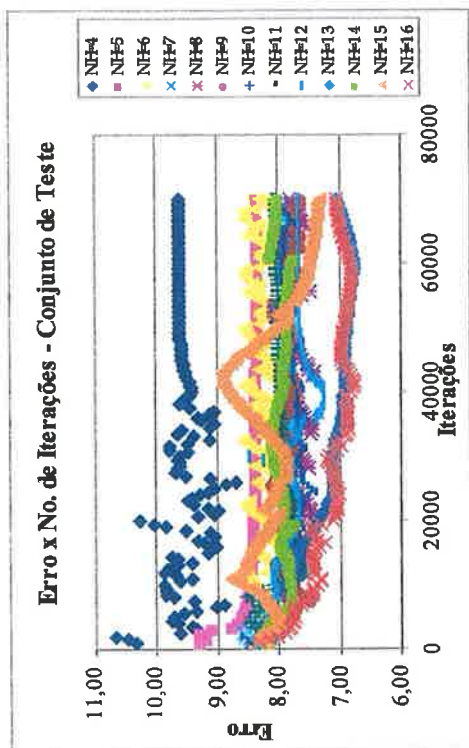


Fig. A.7. – Seleção do Número de Neurônios da REDE 7

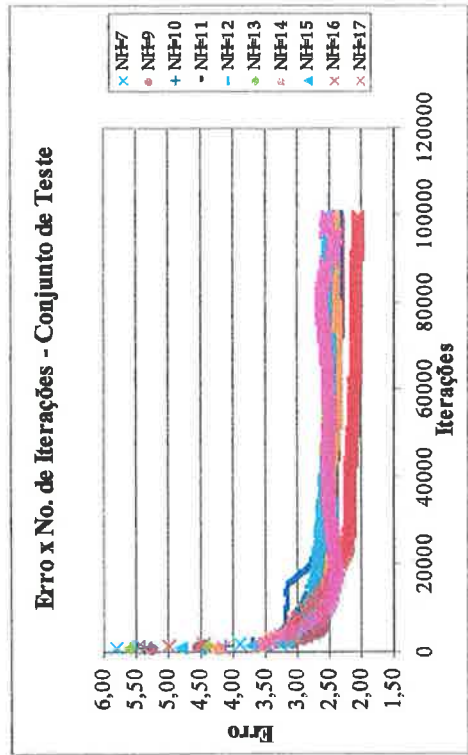
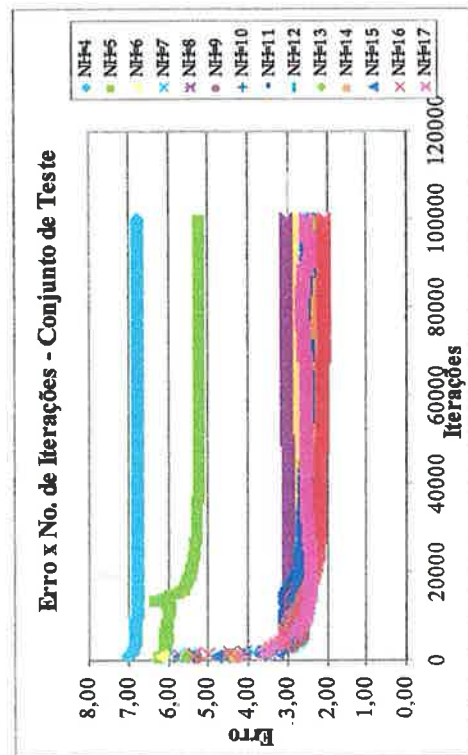


Fig. A.8. – Seleção do Número de Neurônios da REDE 8

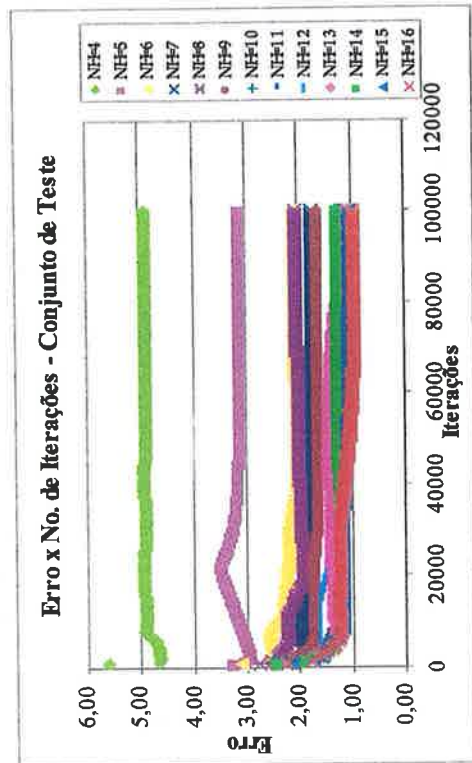
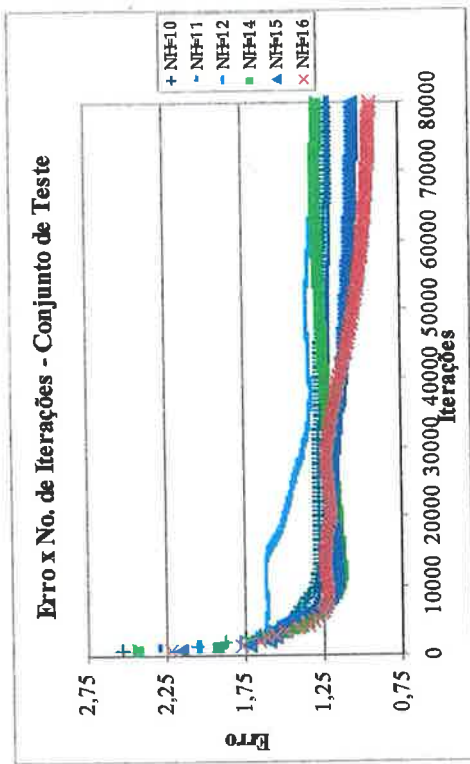


Fig. A.9. – Seleção do Número de Neurônios da REDE 9

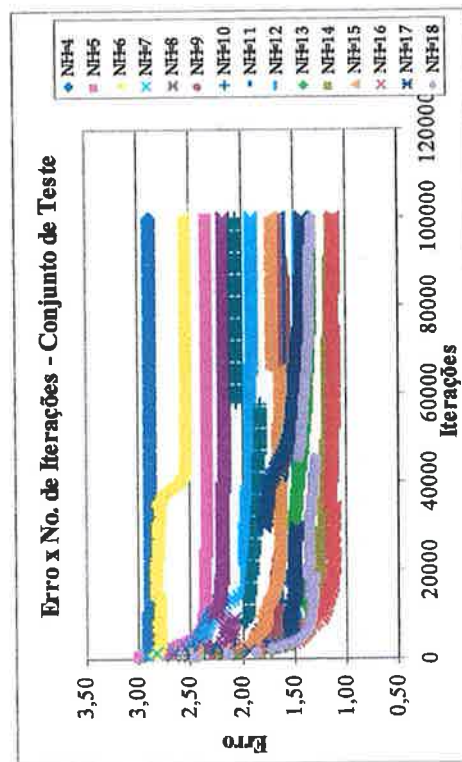
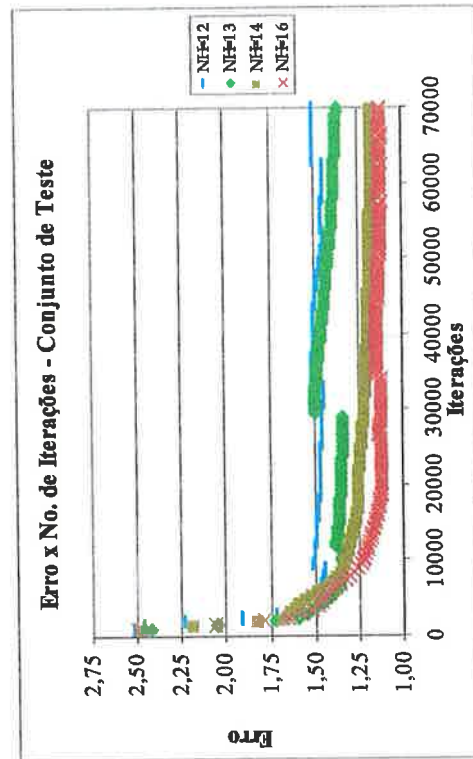


Fig. A.10. – Seleção do Número de Neurônios da REDE 10

ANEXO B - AJUSTE DAS REDES NEURAIS

Neste anexo serão apresentados os gráficos mostrando os resultados dos testes empregados, neste trabalho, para a avaliação do ajuste das redes selecionadas: análise de dispersão dos dados experimentais e calculados, função distância ou distribuição dos erros, histograma dos erros. As unidades das variáveis são arbitrárias.

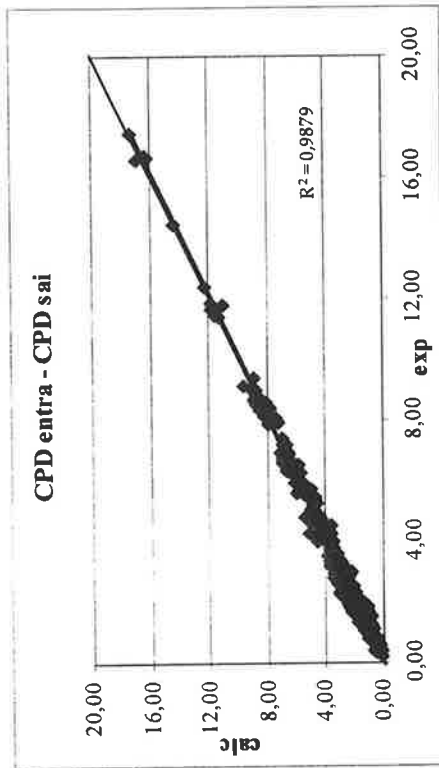
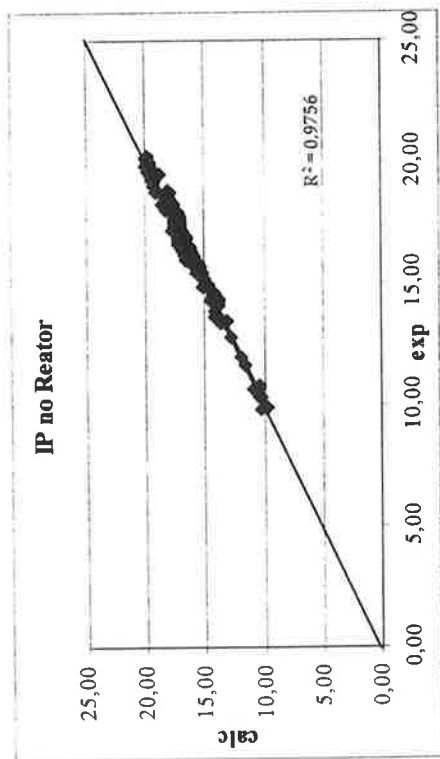


Fig. B.1. – Ajuste da REDE 1 – R-01 para o Conjunto de Teste

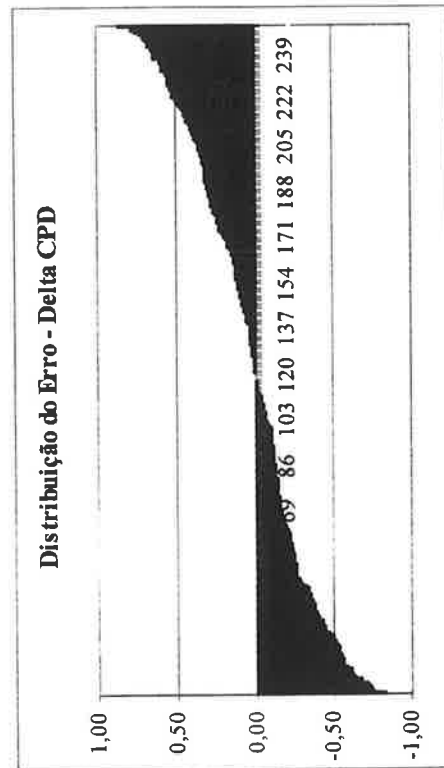
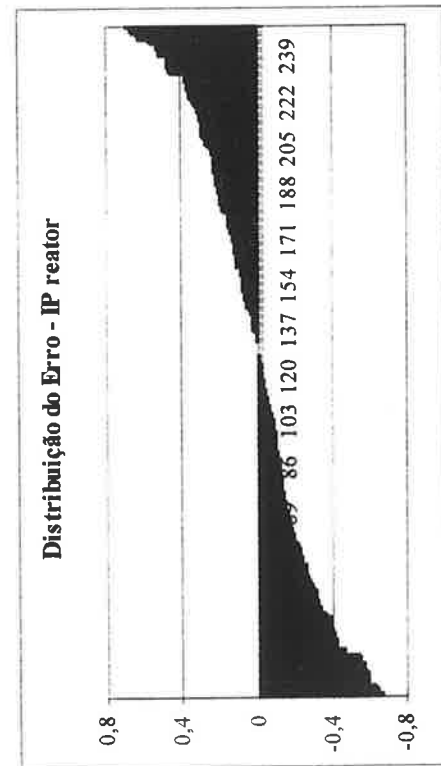


Fig. B.2 – Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 1 – R-01

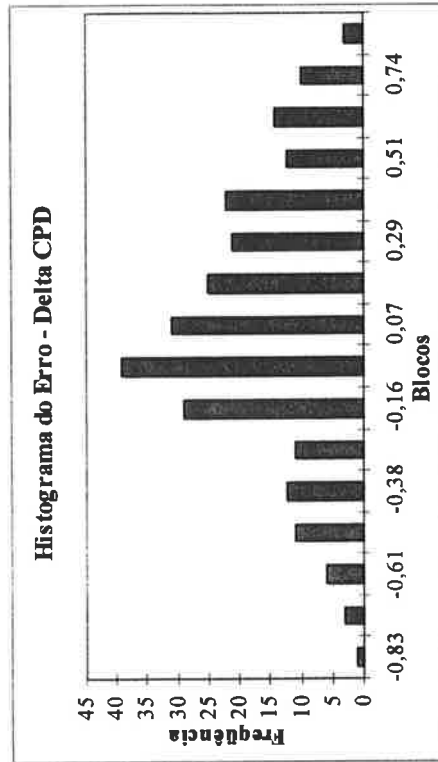
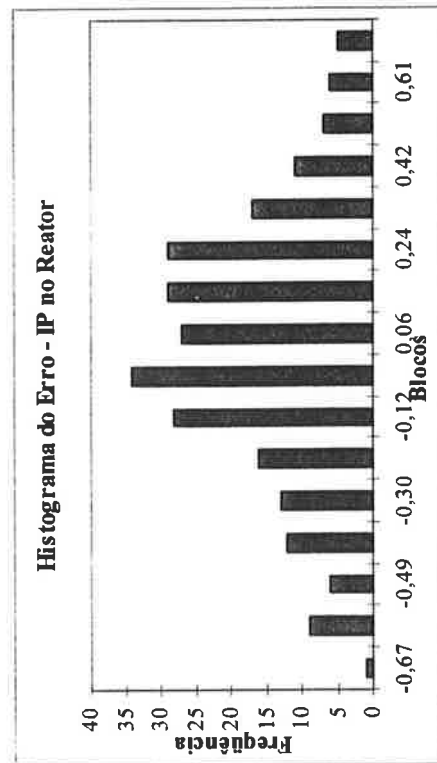


Fig. B.3 – Histograma do Erro – Variáveis de Saída da REDE 1 - R-01

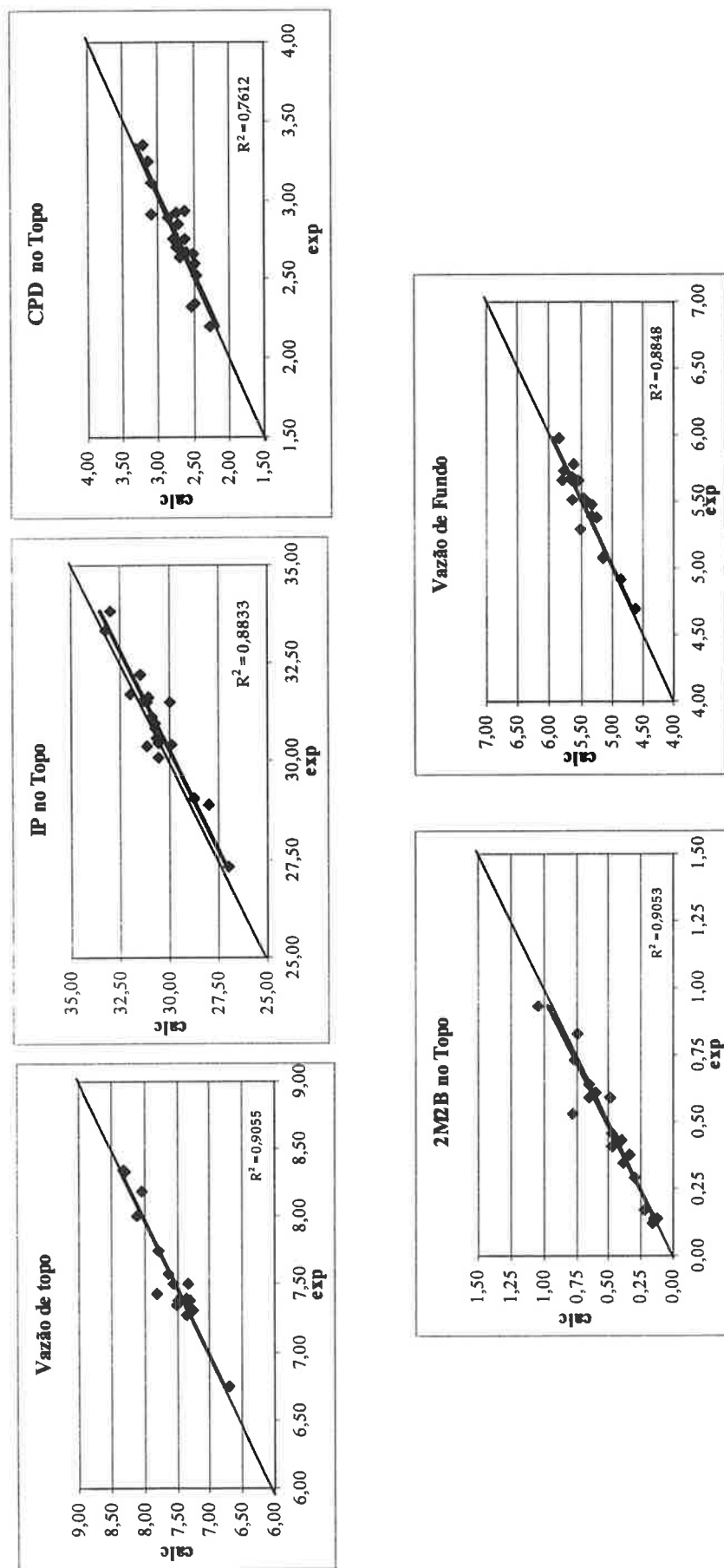


Fig. B.4. – Ajuste da REDE 2 – COL-01 para o Conjunto de Teste

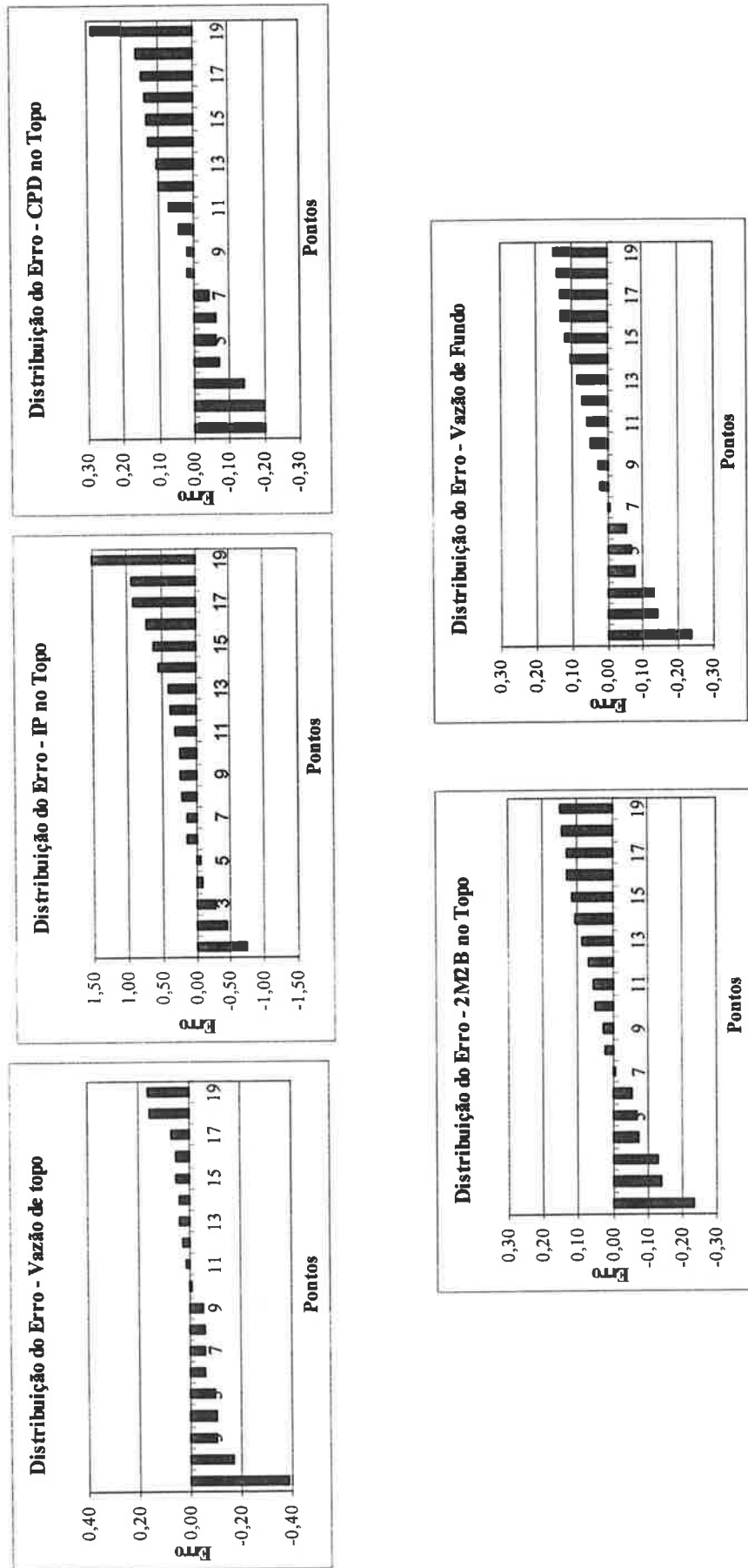


Fig. B.5 – Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 2 – COL-01

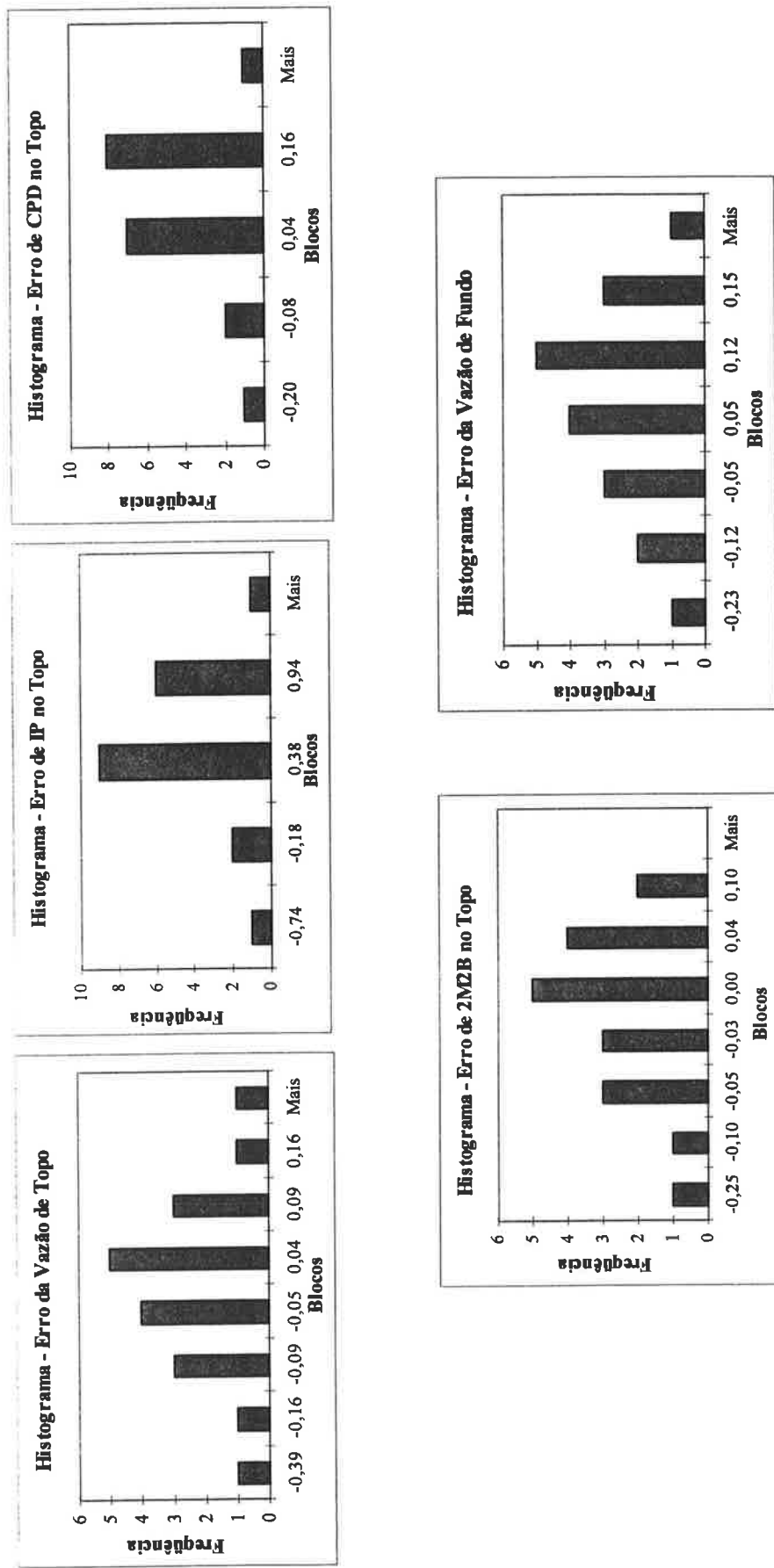


Fig. B.6. — Histograma do Erro — Variáveis de Saída da REDE 2- COL-01

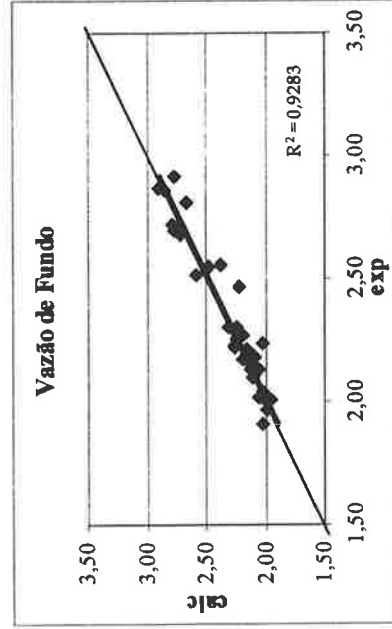
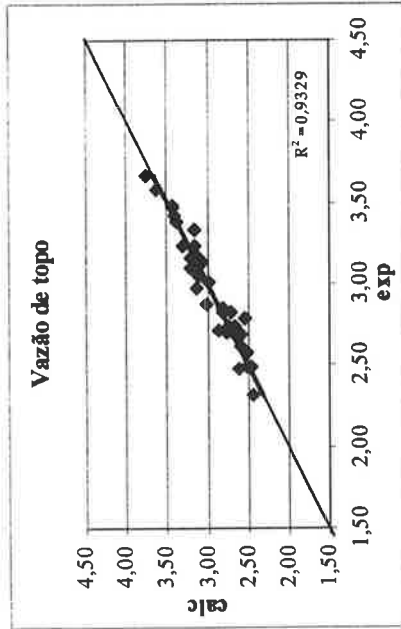
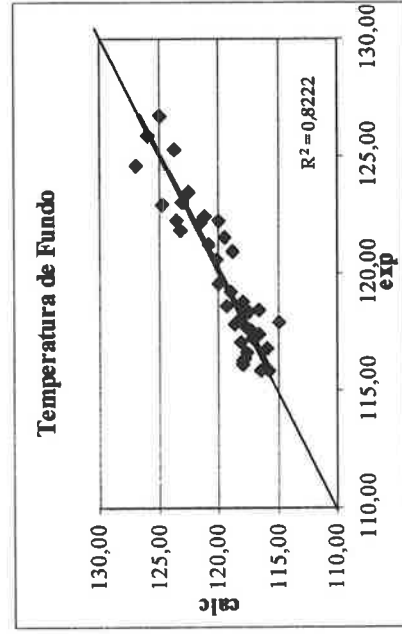
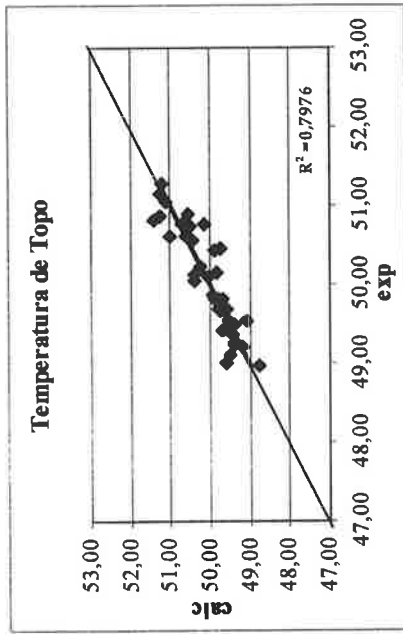


Fig. B.7. – Ajuste da REDE 3 – COL-12 para o Conjunto de Teste

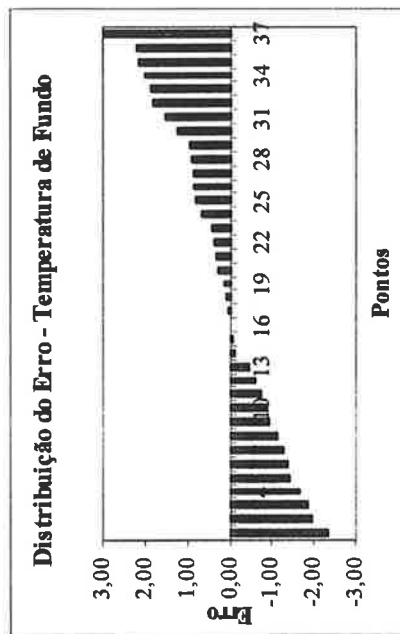
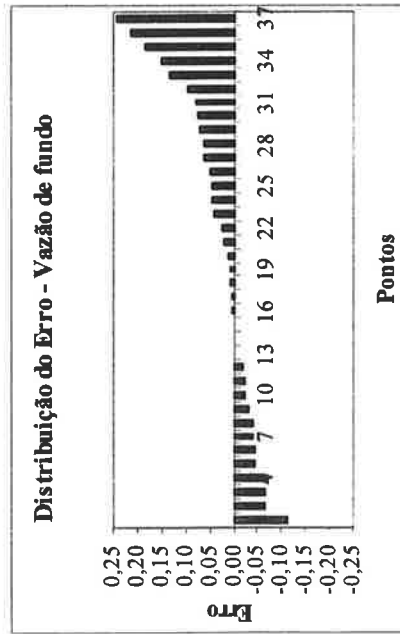
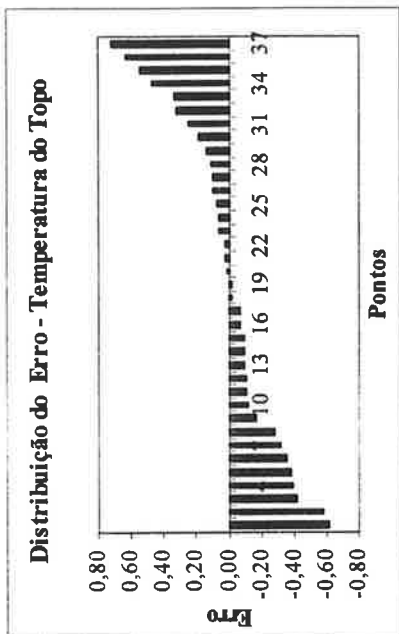
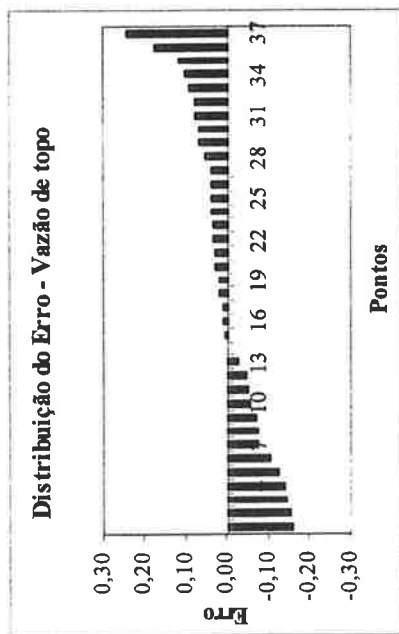


Fig. B.8 – Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 3 – COL-12

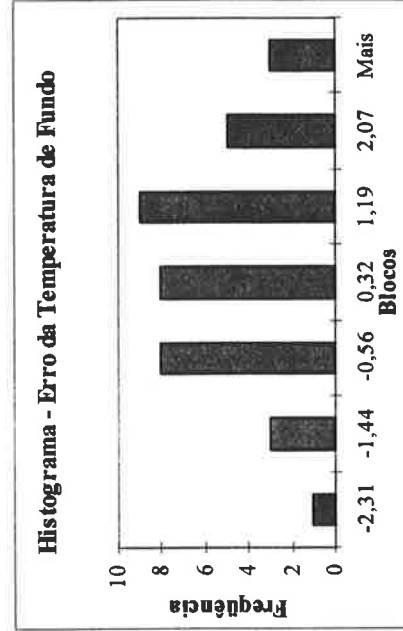
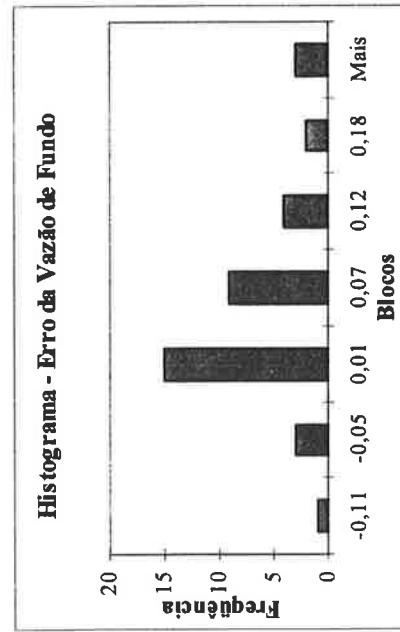
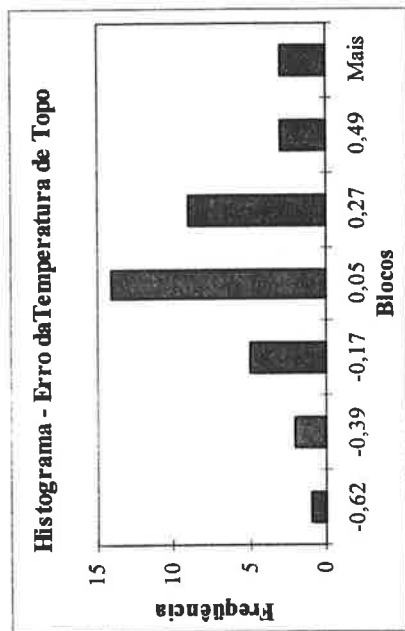
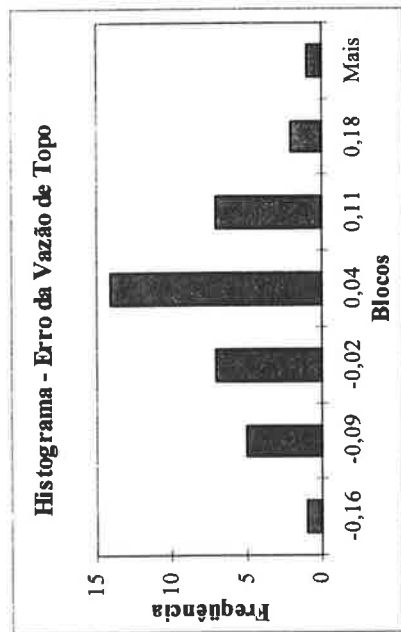


Fig. B.9 – Histograma do Erro – Variáveis de Saída da REDE 3-COL-12

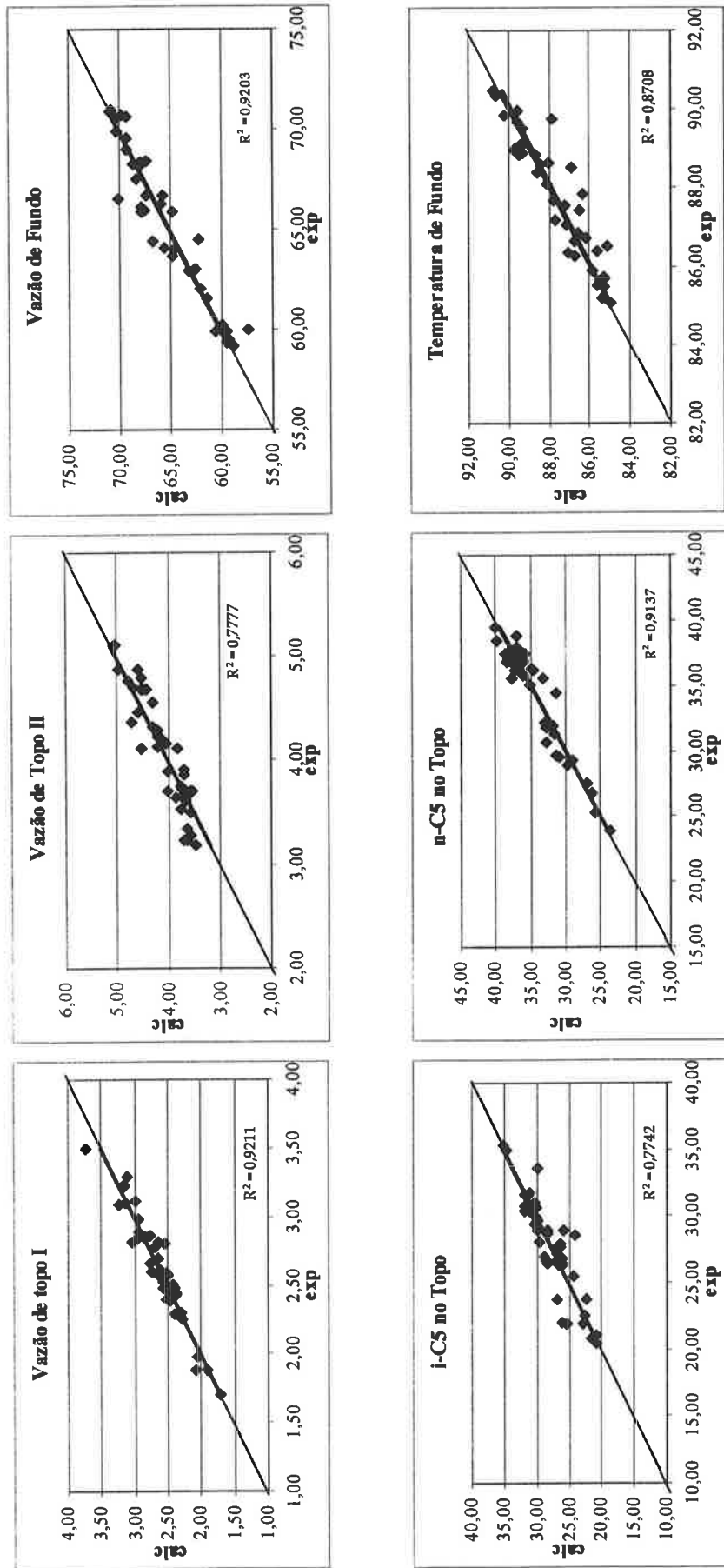


Fig. B.10. – Ajuste da REDE 4 – COL-02 para o Conjunto de Teste

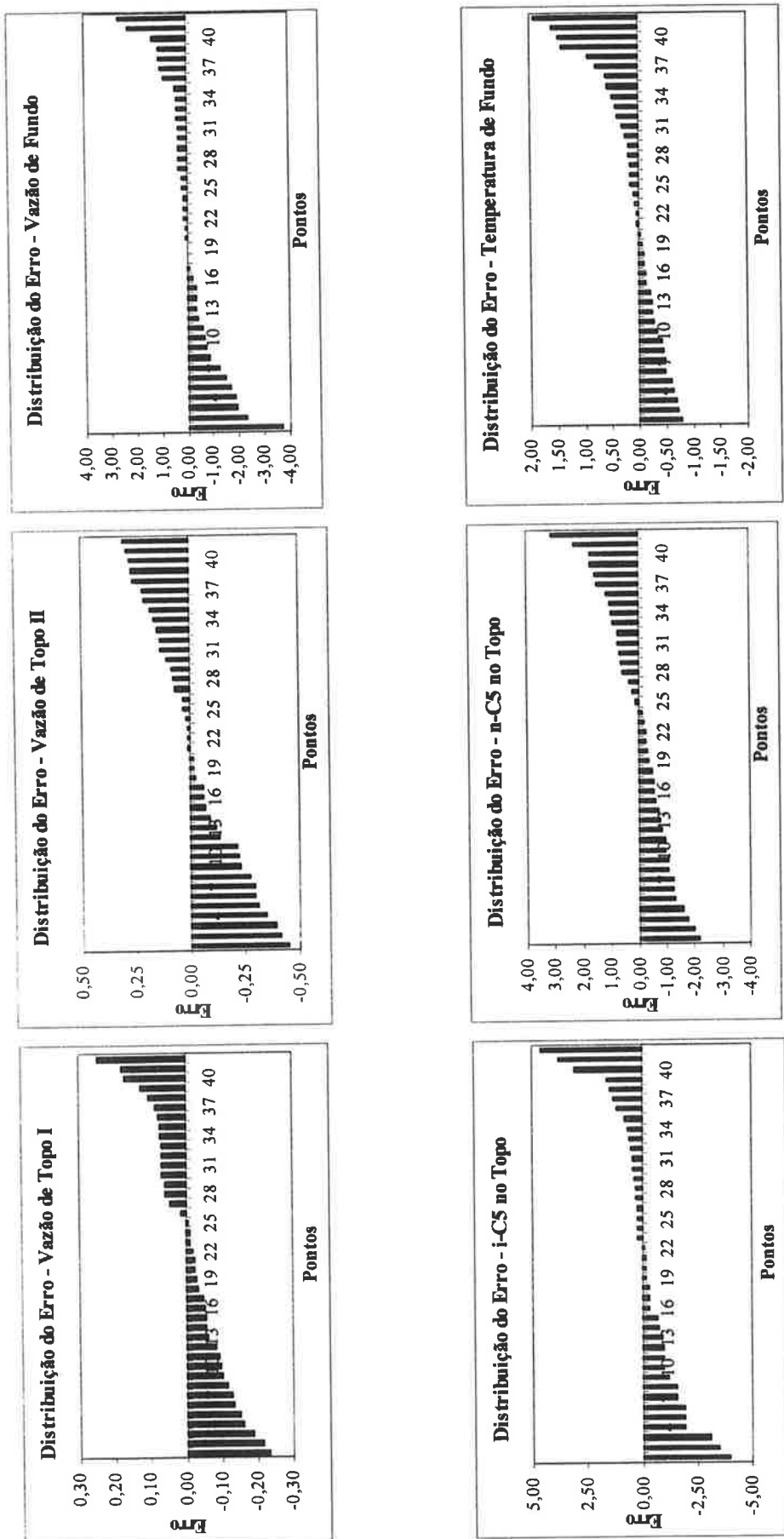


Fig. B.11 – Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 4 – COL-02

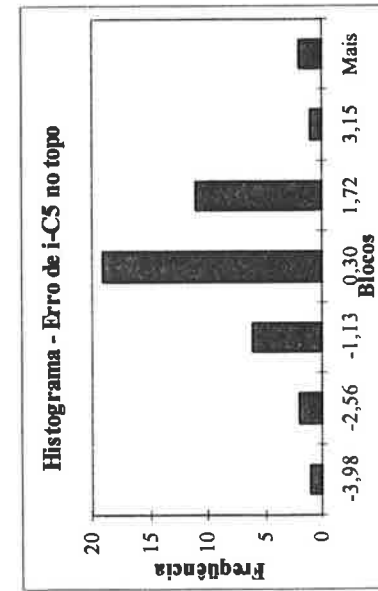
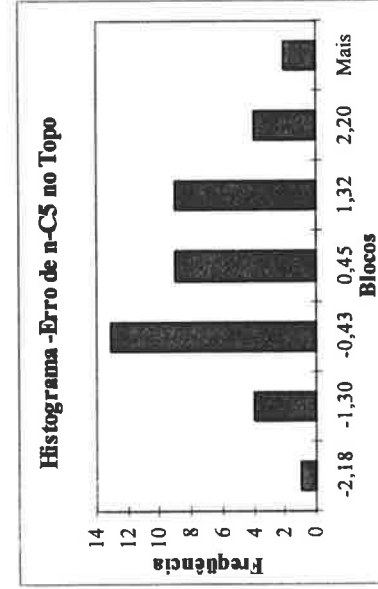
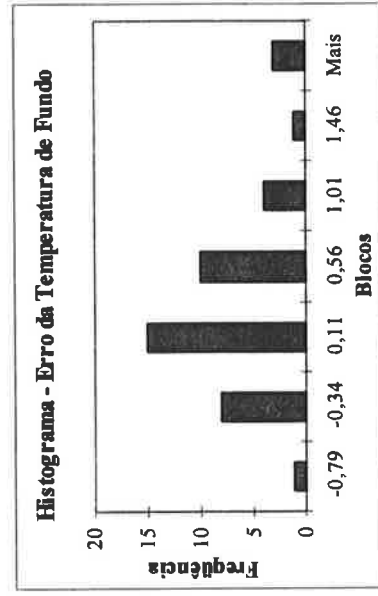
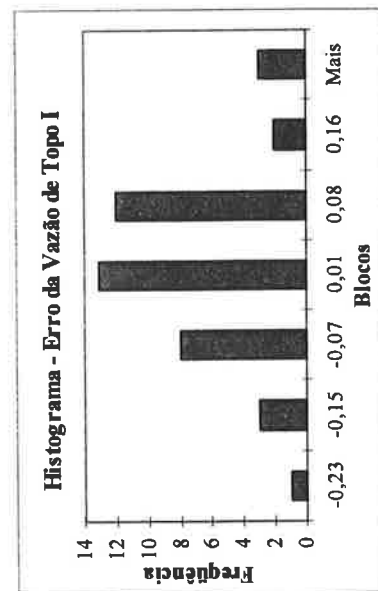
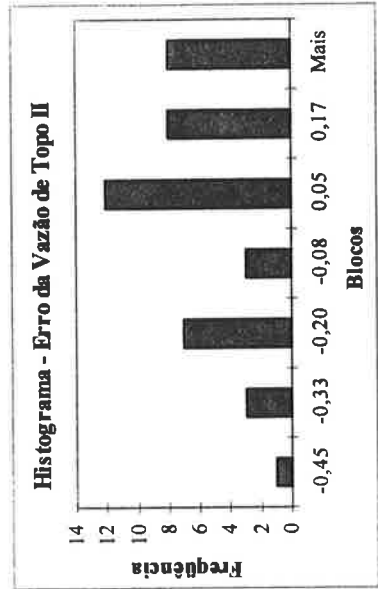
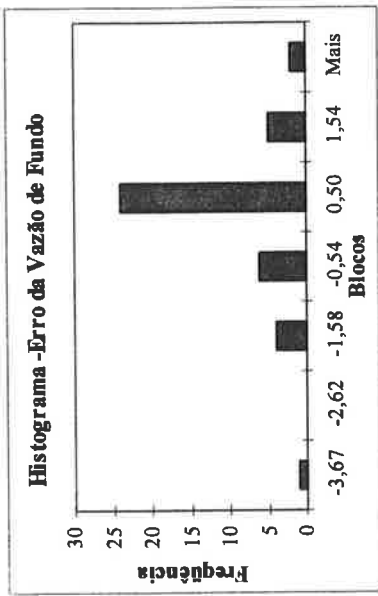


Fig. B.12 – Histograma do Erro – Variáveis de Saída da REDE 4 - COL-02

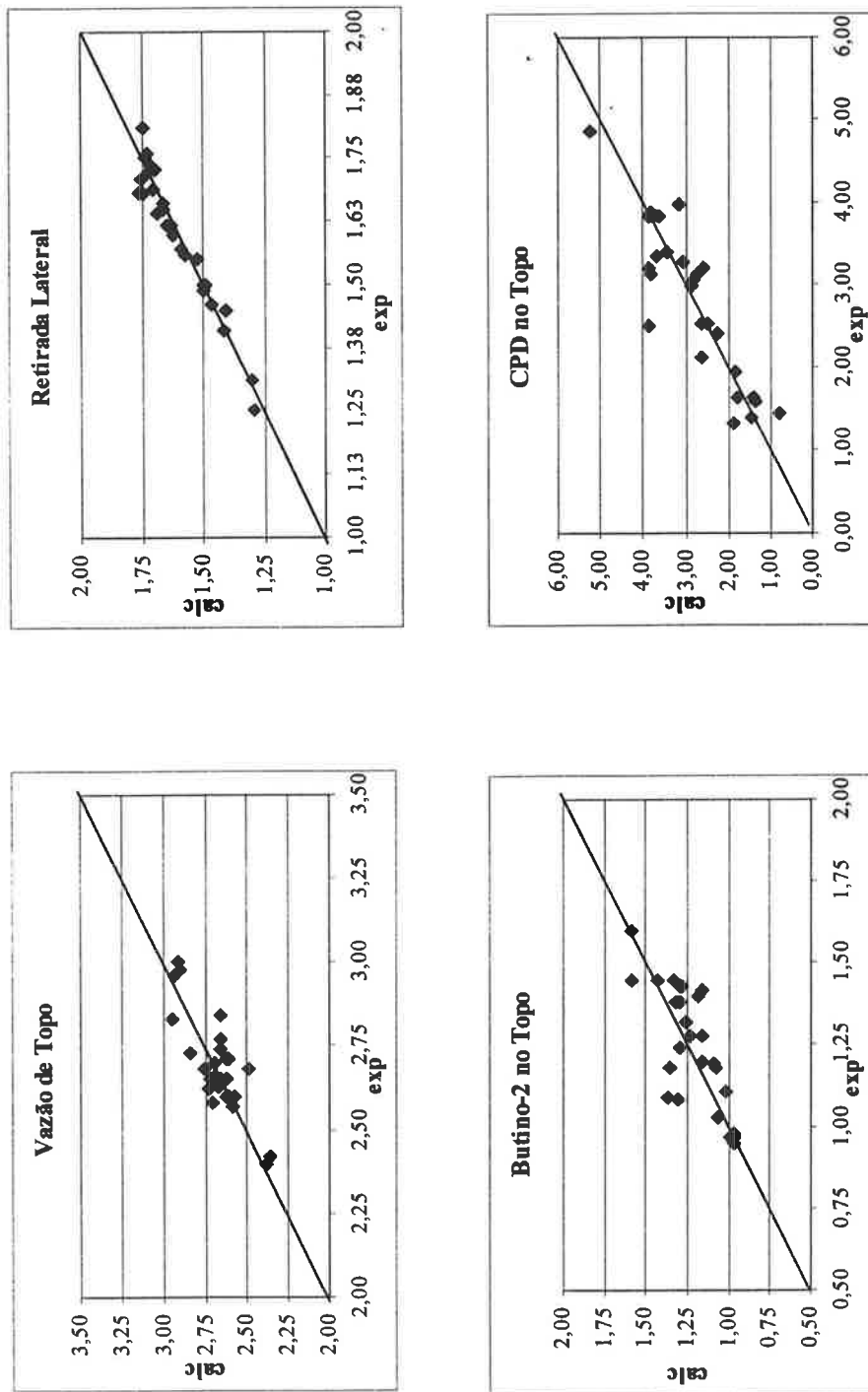


Fig. B.13. – Ajuste da REDE 5 – COL-03 e COL-04 para o Conjunto de Teste

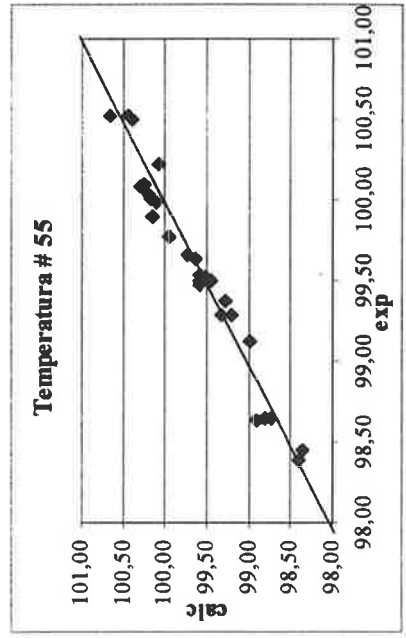
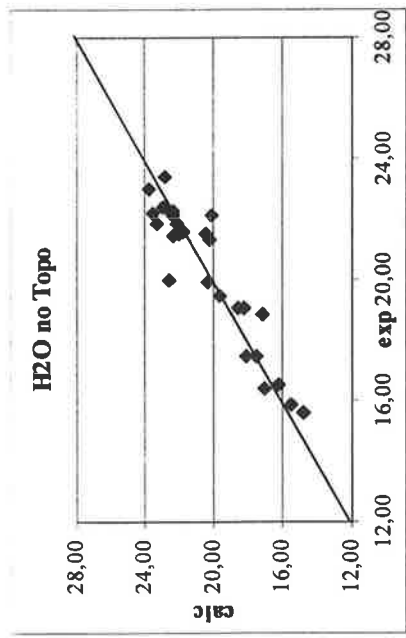
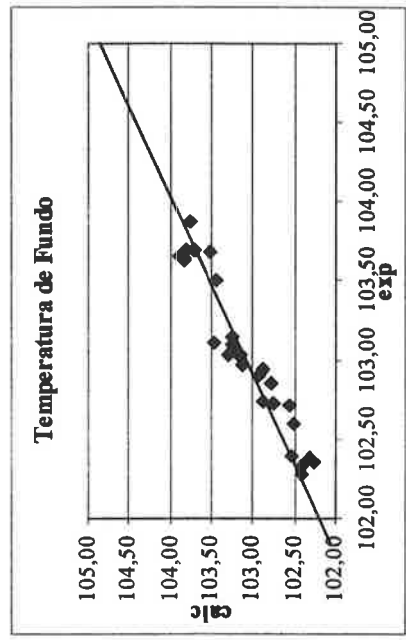
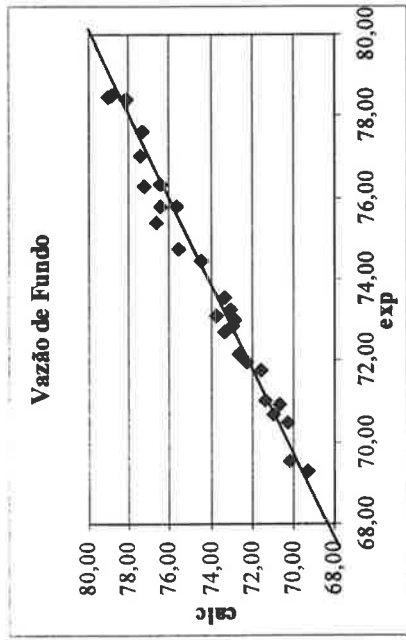


Fig. B.13. (cont.) – Ajuste da REDE 5 – COL-03 e COL-04 para o Conjunto de Teste

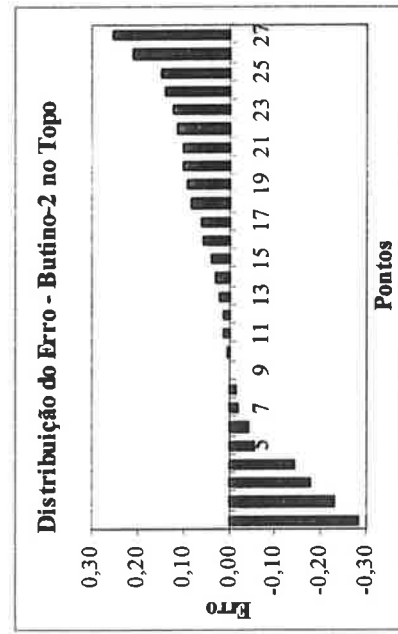
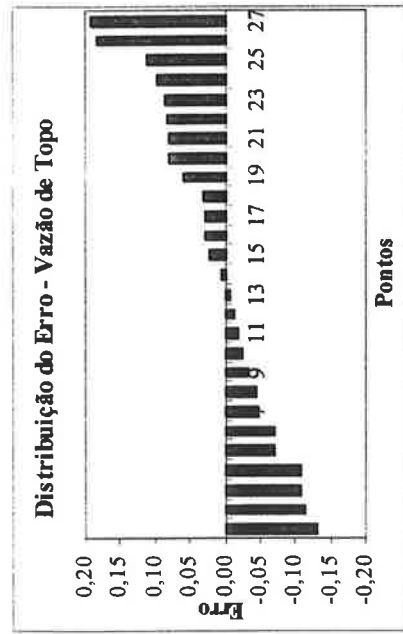
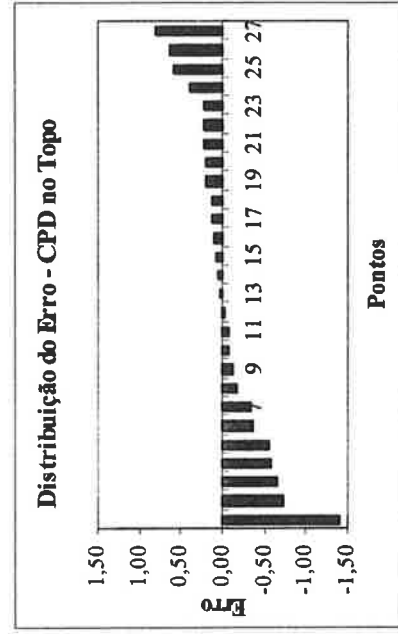
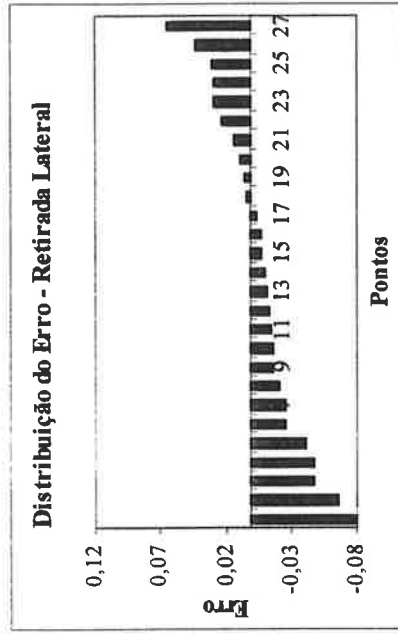


Fig. B.14 – Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 5 – COL-03 e COL-04

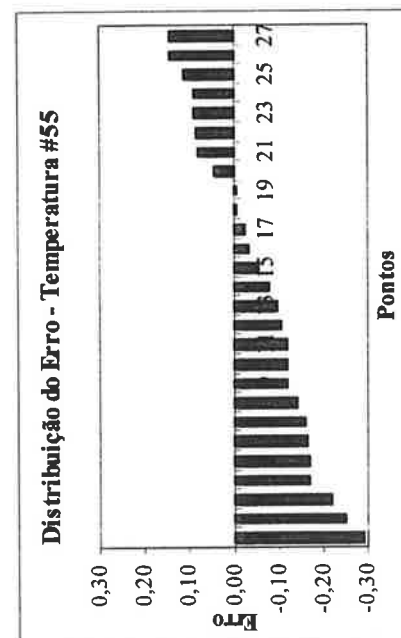
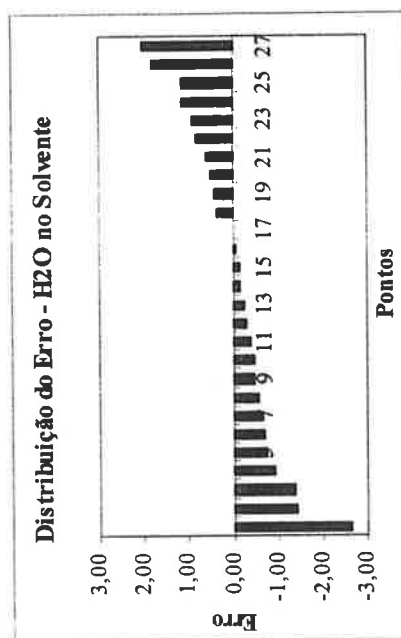
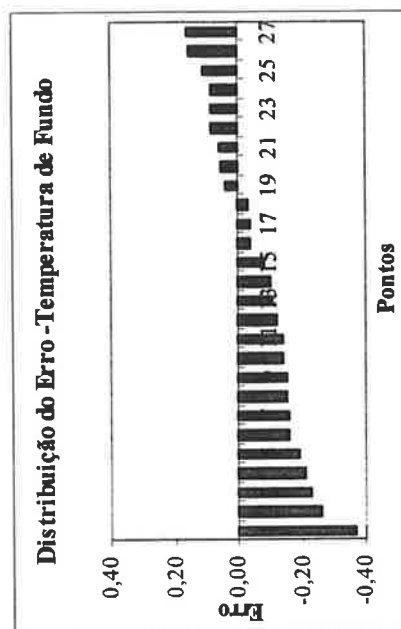
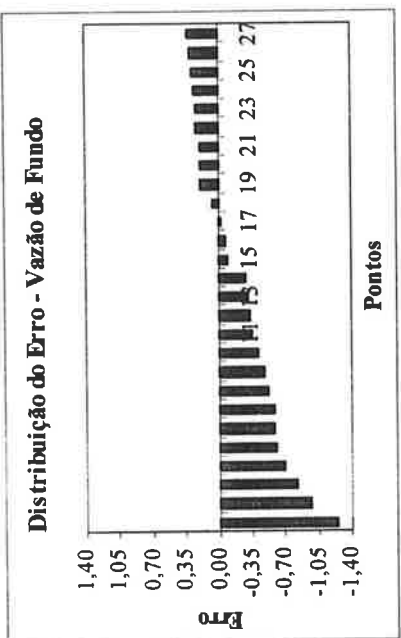


Fig. B.14. (cont.) – Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 5 – COL-03 e COL-04

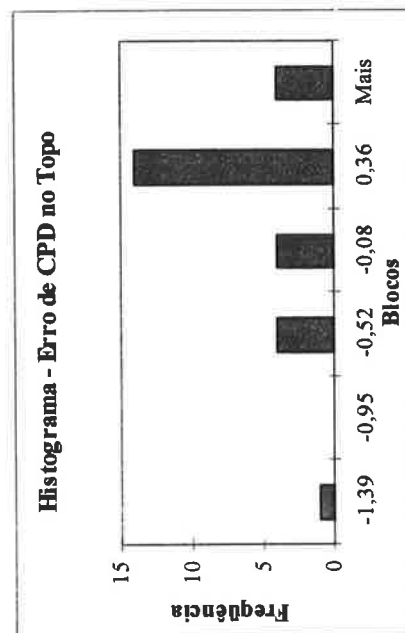
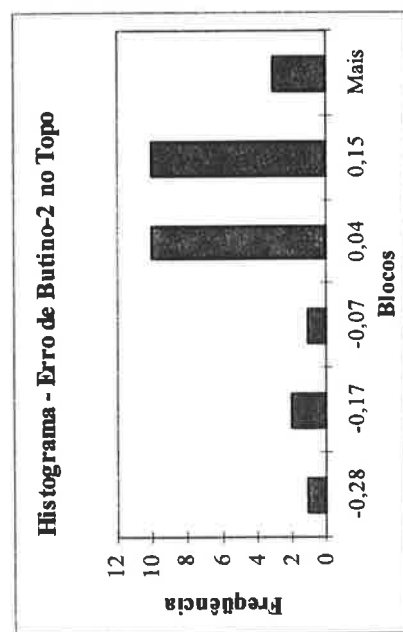
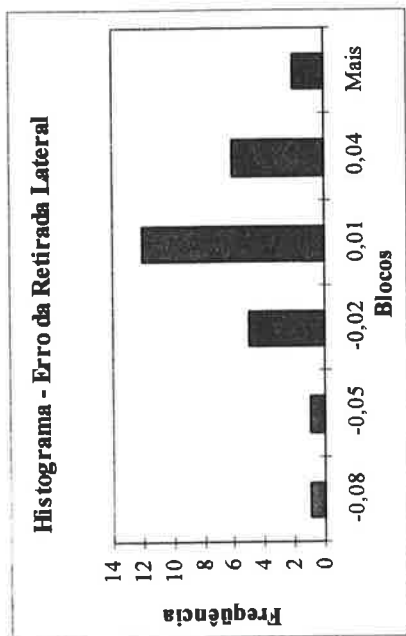
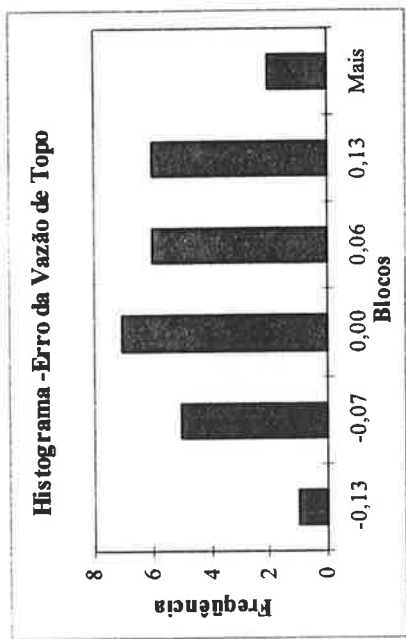


Fig. B.15 – Histograma do Erro – Variáveis de Saída da REDE 5- COL-03 e COL-04

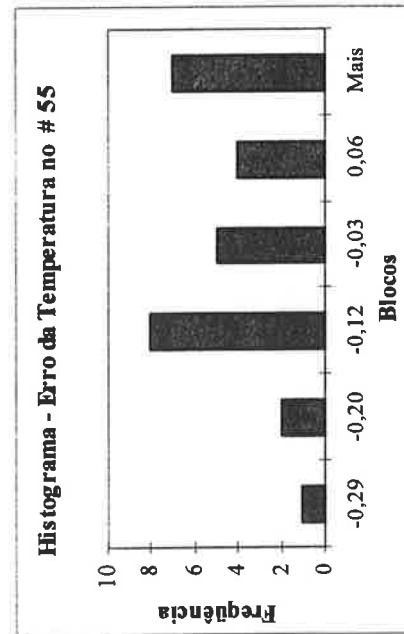
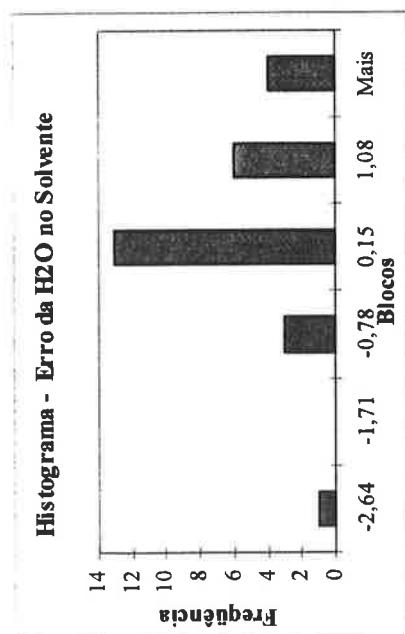
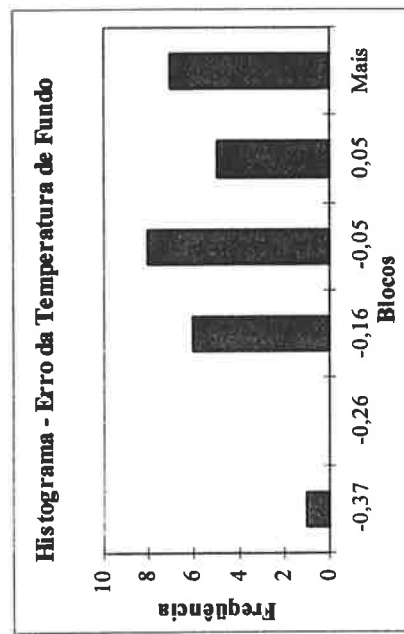
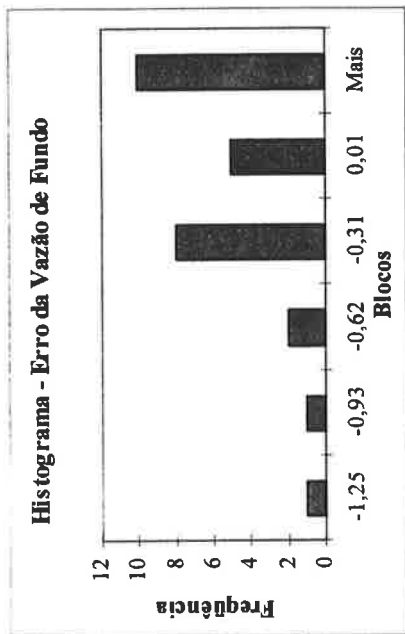


Fig. B.15. (cont.) – Histograma do Erro -- Variáveis de Saída da REDE 5- COL-03 e COL-04

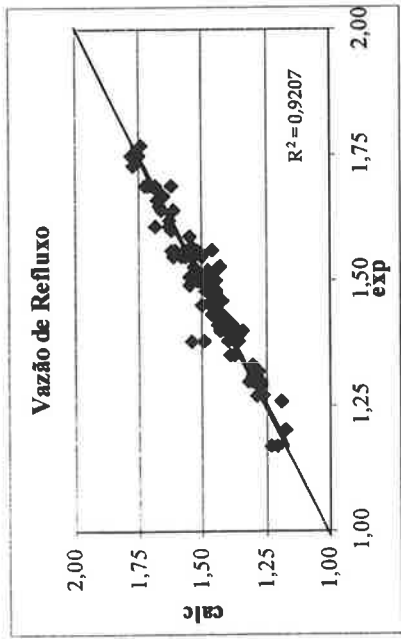
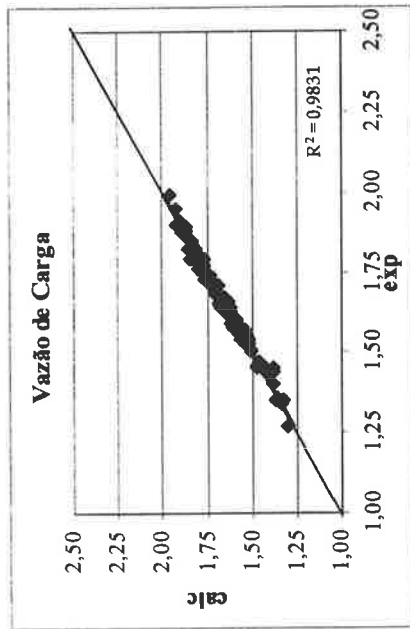


Fig. B.16 – Ajuste da REDE 6 – COL-05 para o Conjunto de Teste

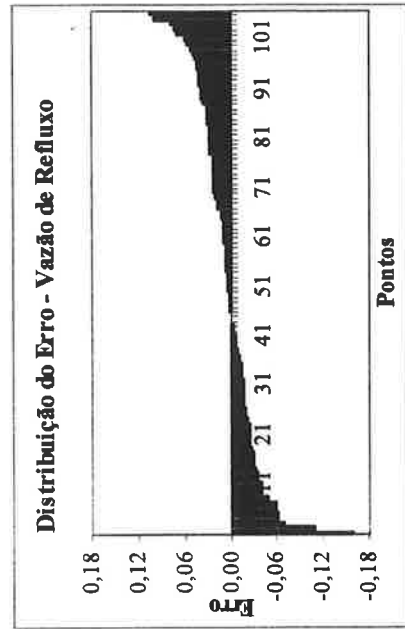
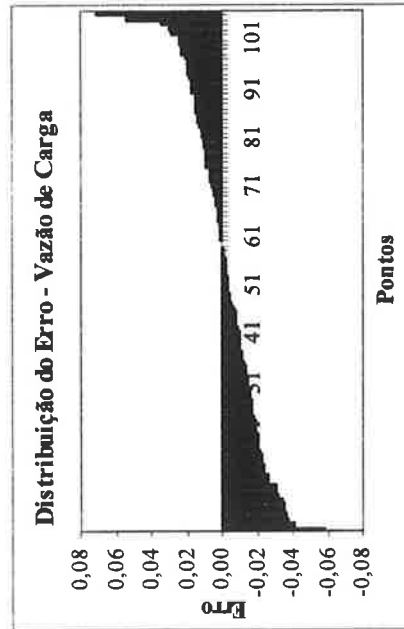


Fig. B.17 – Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 6 – COL-05

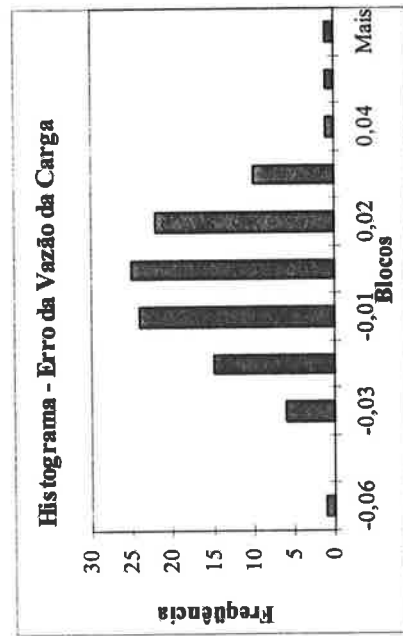
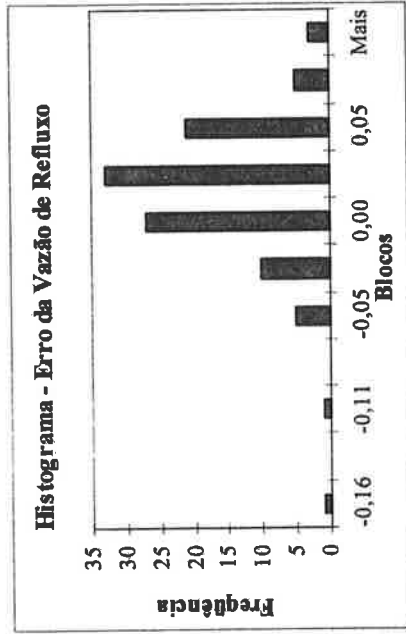


Fig. B.18 – Histograma do Erro – Variáveis de Saída da REDE 6- COL-05

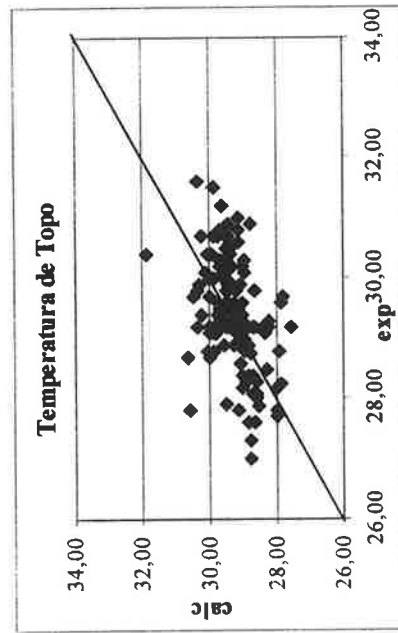
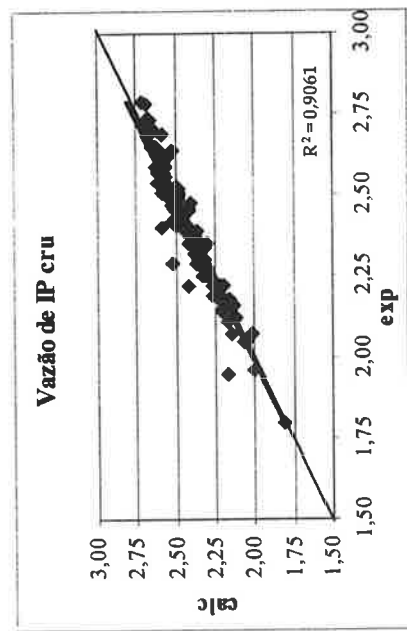
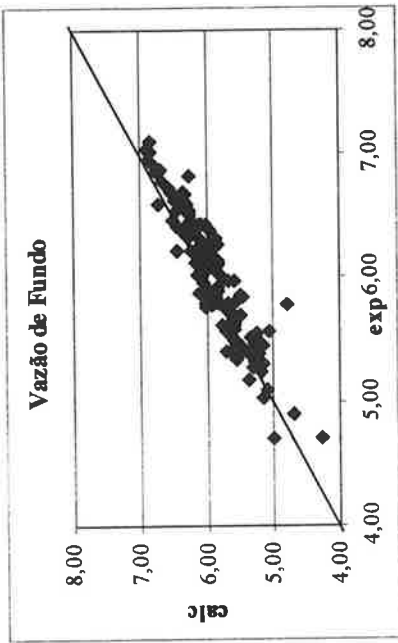


Fig. B.19 – Ajuste da REDE 7 – COL-06 e COL-07 para o Conjunto de Teste

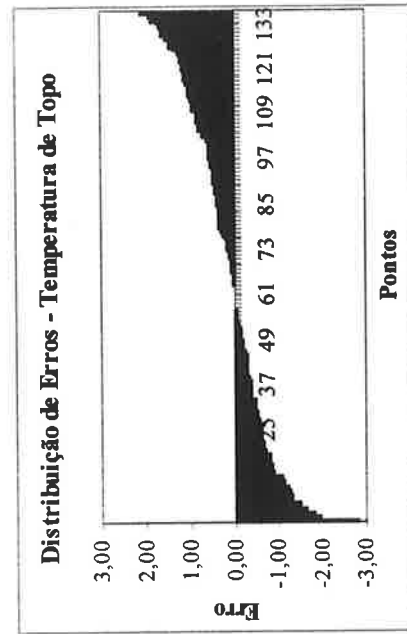
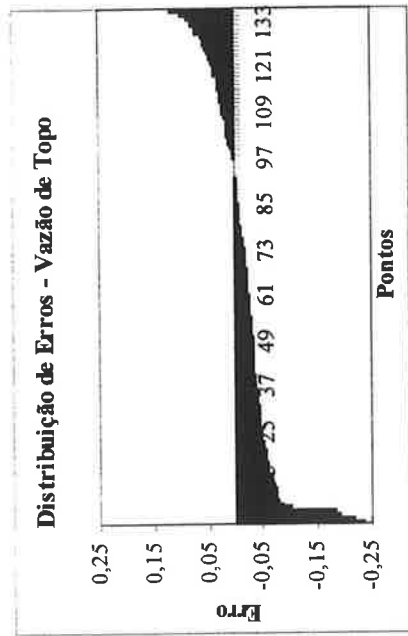
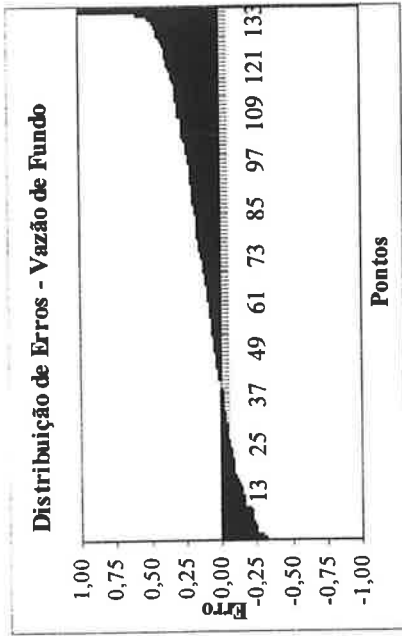


Fig. B.20 – Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 7 – COL-06 e COL-07

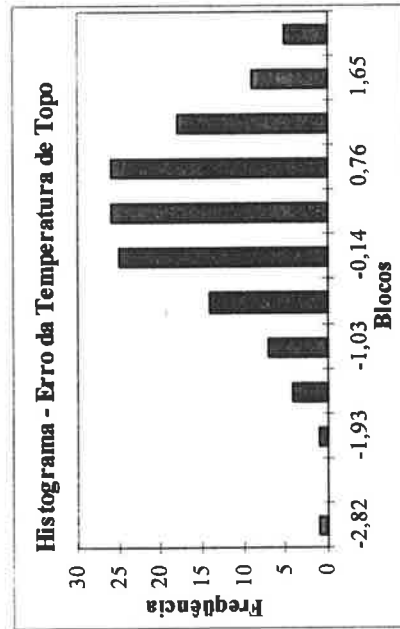
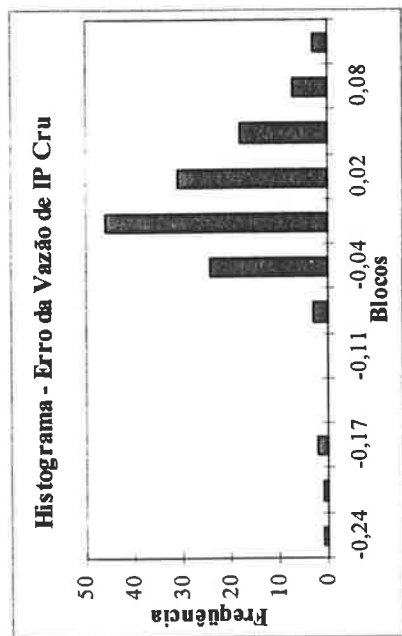
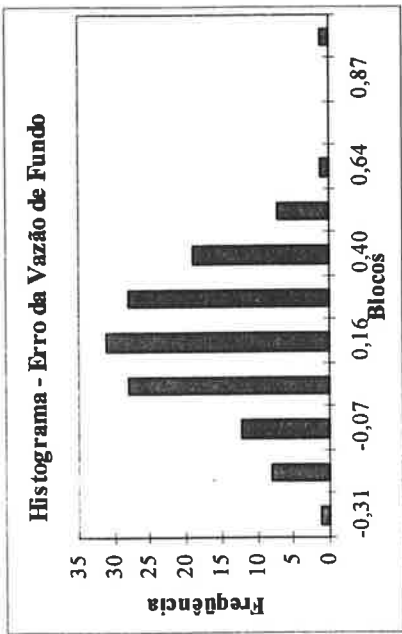


Fig. B.21 – Histograma do Erro – Variáveis de Saída da REDE 7- COL-06 e COL-07

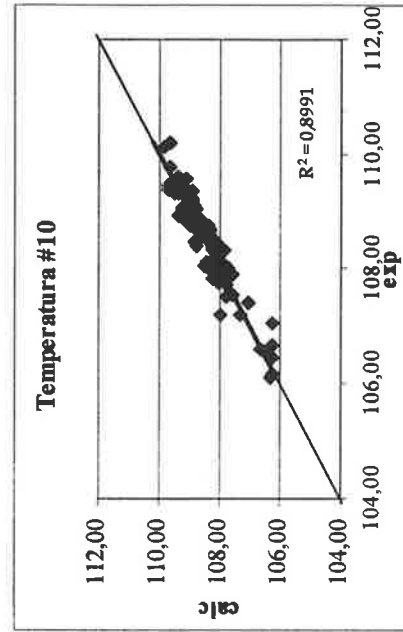
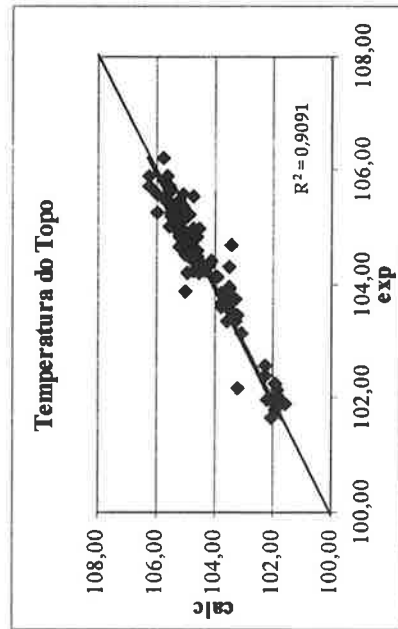
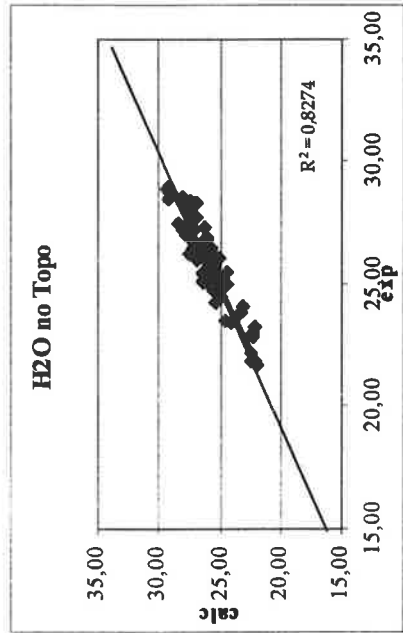
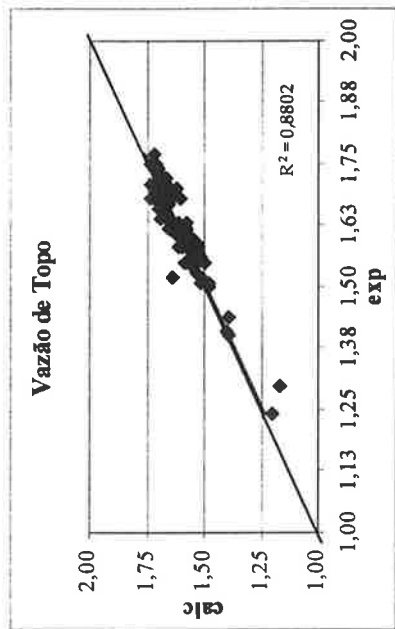


Fig. B.22 – Ajuste da REDE 8 – COL-10 para o Conjunto de Teste

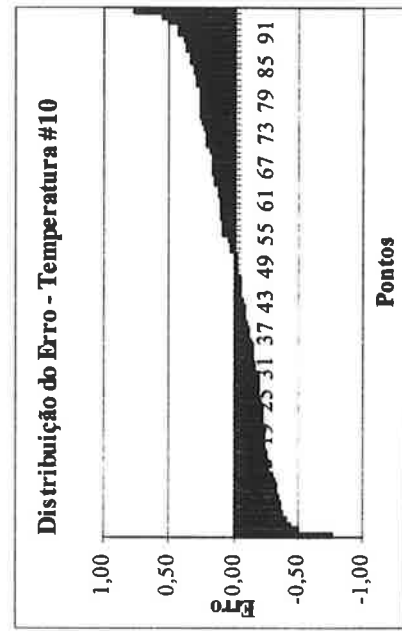
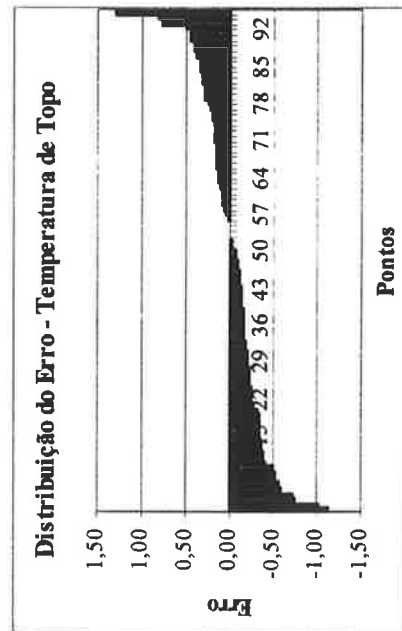
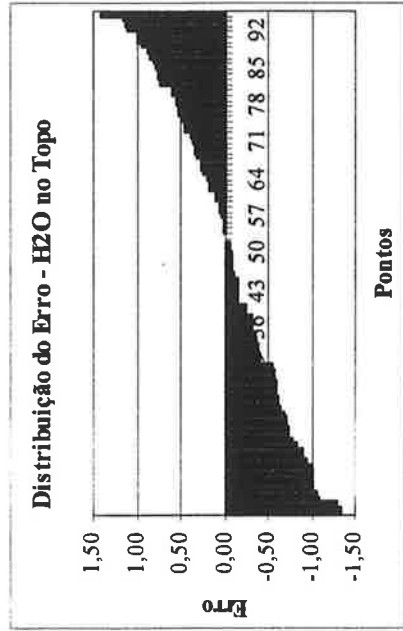
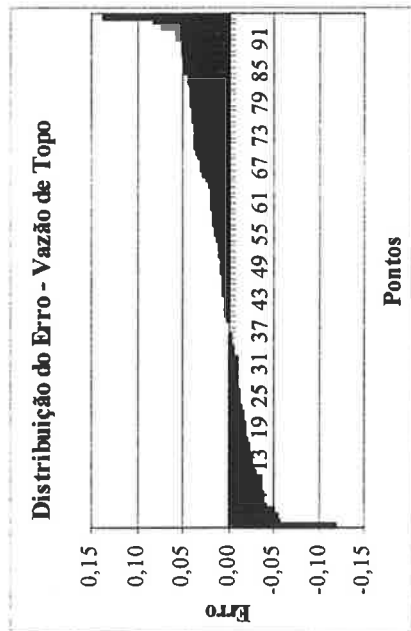


Fig. B.23 – Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 8 – COL-10

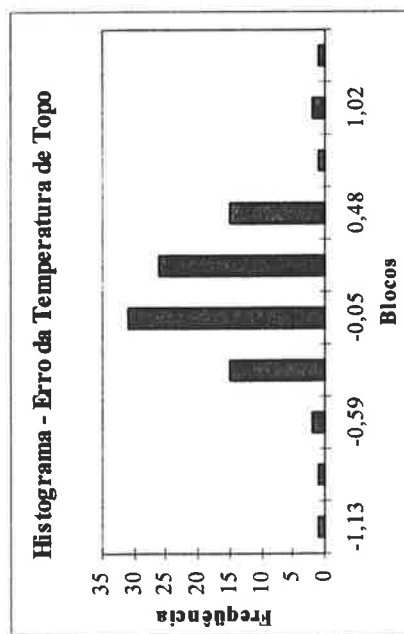
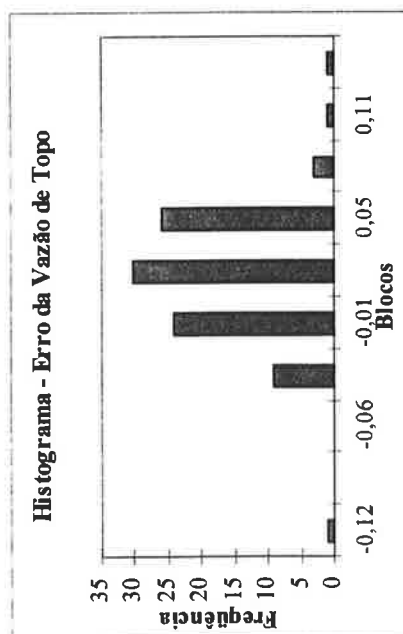
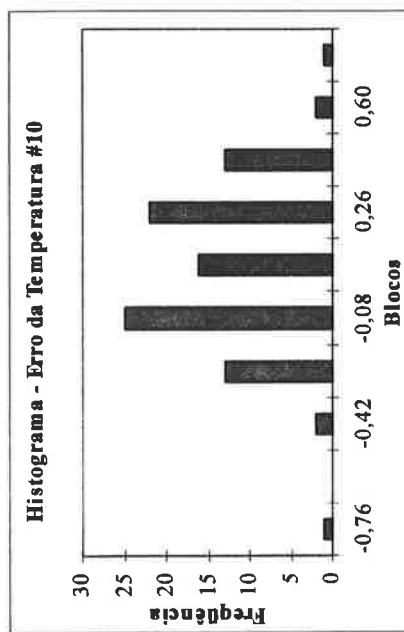
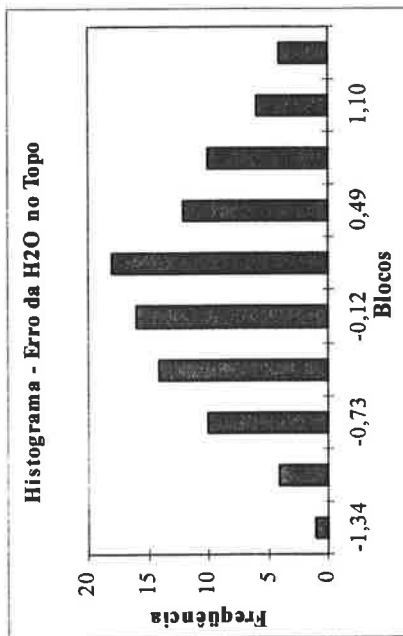


Fig. B.24 – Histograma do Erro – Variáveis de Saída da REDE 8 - COL-10

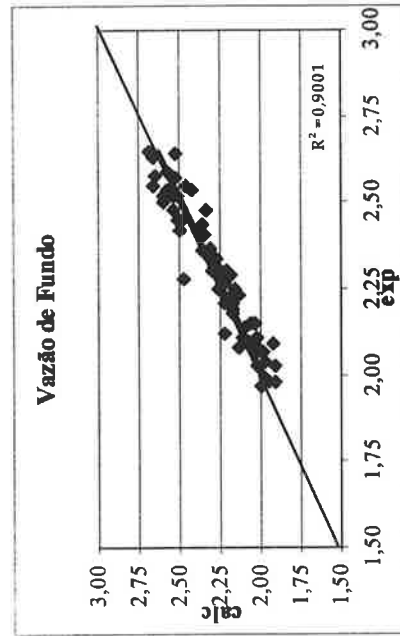
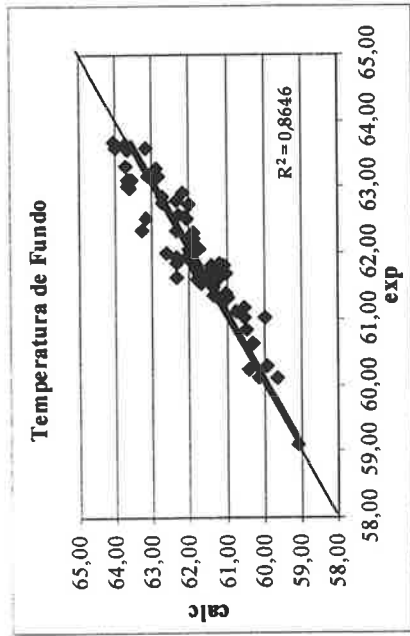
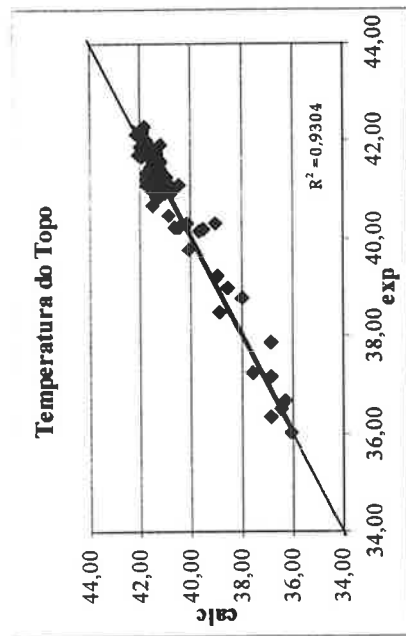


Fig. B.25 – Ajuste da REDE 9 – COL-08 para o Conjunto de Teste

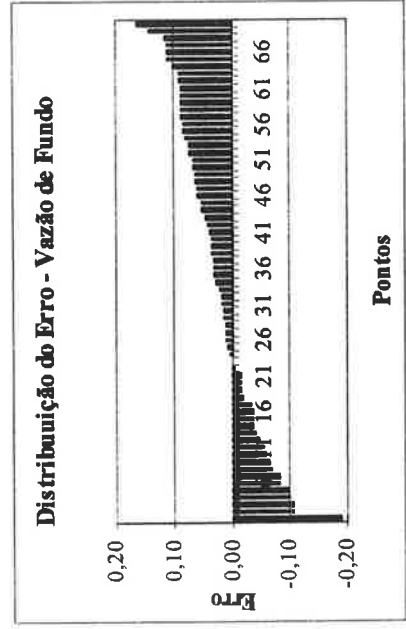
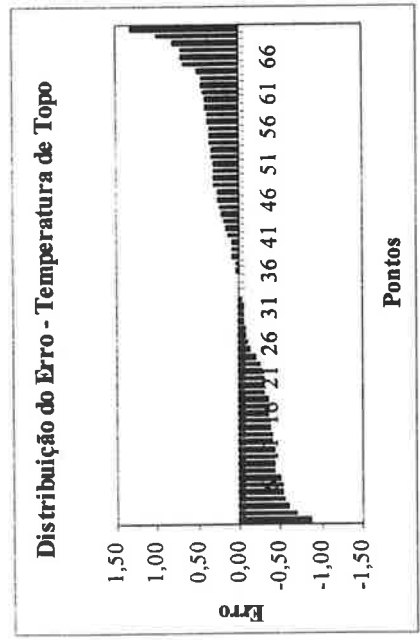
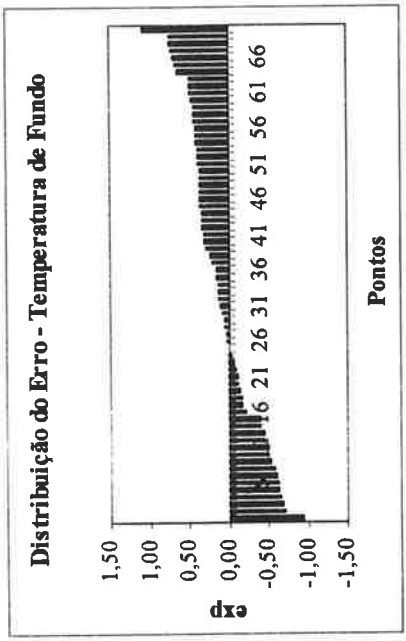


Fig. B.26 – Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 9 – COL-08

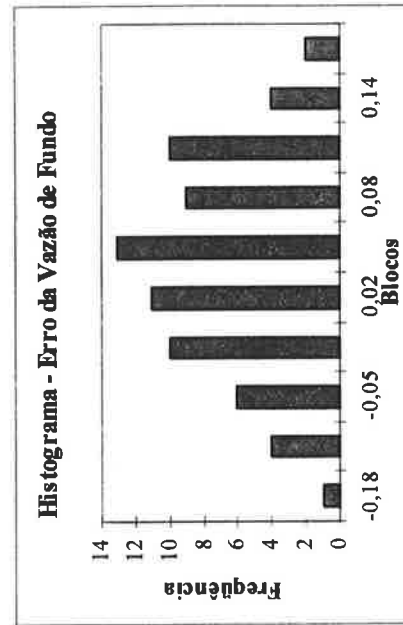
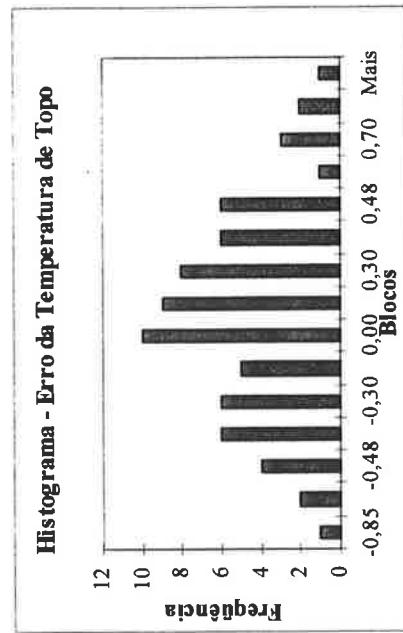
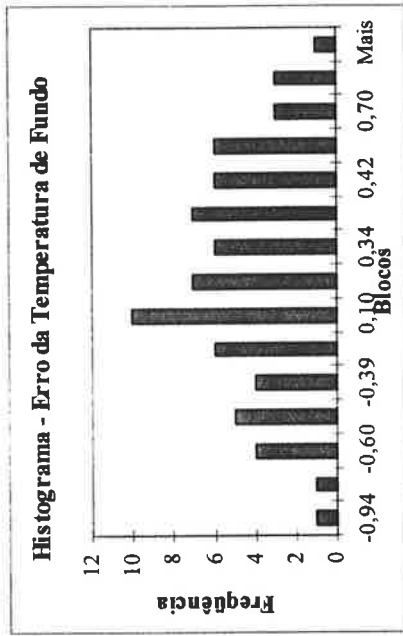


Fig. B.27 – Histograma do Erro – Variáveis de Saída da REDE 9- COL-08

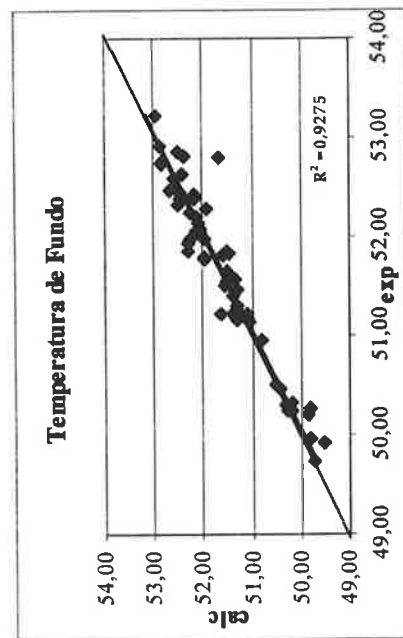
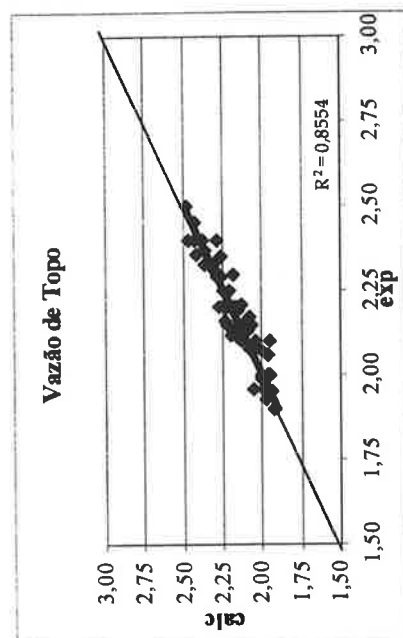
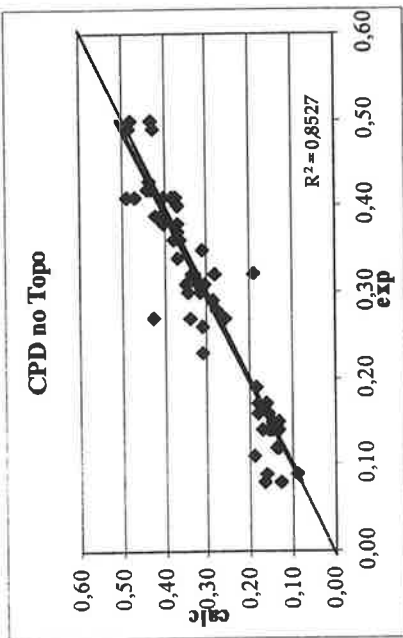


Fig. B.28 – Ajuste da REDE 10 – COL-09 para o Conjunto de Teste

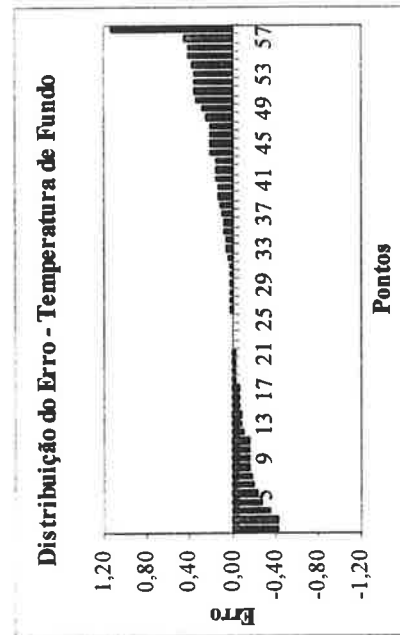
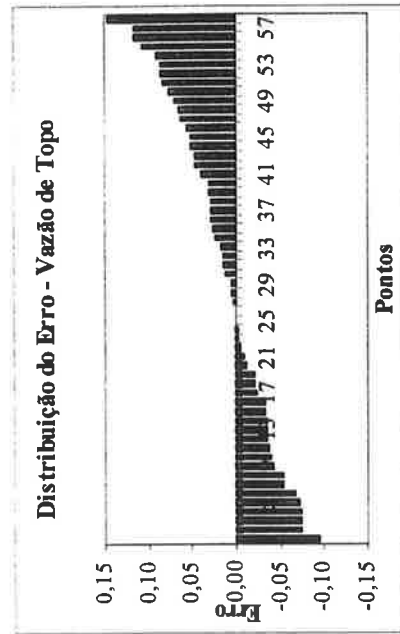
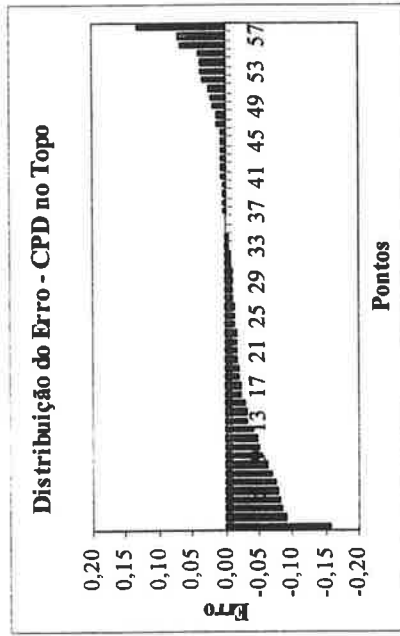


Fig. B.29 – Função Distância – Distribuição do Erro para as Variáveis de Saída da REDE 10 – COL-09

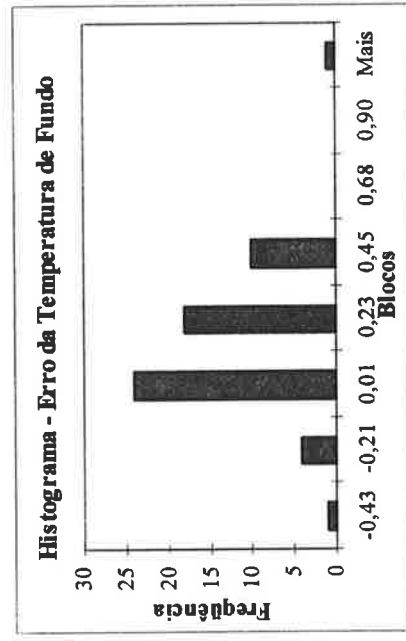
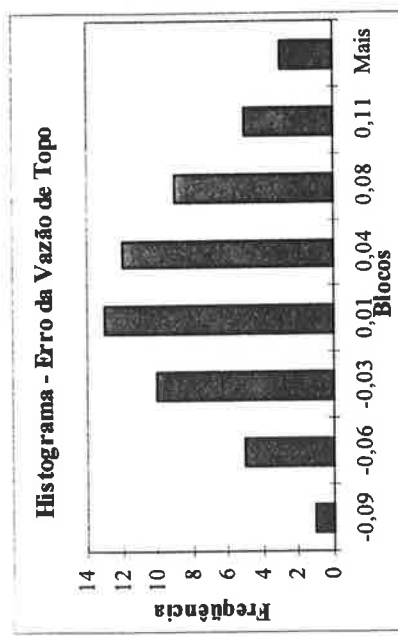
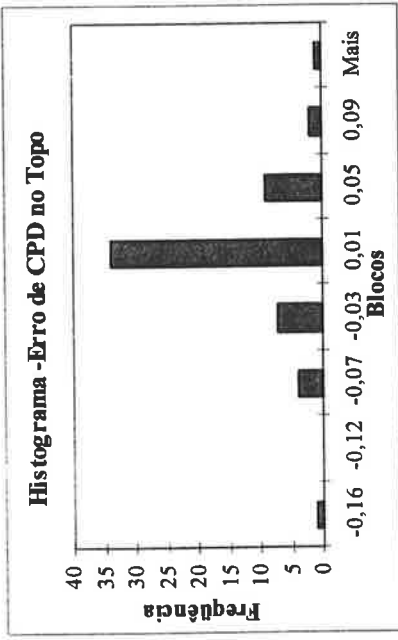


Fig. B.30 – Histograma do Erro – Variáveis de Saída da REDE 10- COL-09

ANEXO C – SIMULAÇÃO DO PROCESSO – TABELAS

As tabelas a seguir mostram as variáveis de entrada os resultados obtidos na simulação do processo para seis diferentes condições operacionais: Caso 1 a Caso 8.

As unidades das variáveis são arbitrárias.

Tabela C.1. – Valores das Variáveis de Entrada para Simulação

Variáveis	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Caso 6	Caso 7	Caso 8
x1(1)	14,02	13,99	14,50	12,49	12,50	9,50	11,96	11,96
x1(2)	14,81	20,72	18,38	20,20	20,60	18,72	19,34	19,34
x1(3)	4,35	14,90	5,24	4,63	9,20	18,77	18,66	18,66
x1(4)	71,17	59,80	68,16	70,06	73,05	65,97	60,09	60,09
x2(1)	13,00	13,50	12,50	13,50	13,00	12,80	12,50	12,51
x2(2)	65,63	62,61	65,19	65,70	64,80	63,59	63,17	63,68
x2(3)	17,44	17,33	18,46	18,19	16,88	18,08	18,79	18,49
x2(4)	3,91	4,79	0,83	4,26	3,93	3,37	3,62	5,78
x2(5)	50,15	44,24	46,18	48,12	44,53	39,07	34,80	36,36
x2(6)	11,12	9,40	10,01	10,01	9,16	8,99	8,65	8,82
x3(1)	5,22	4,89	5,92	5,23	4,88	6,11	5,80	5,55
x3(2)	2,80	2,80	2,30	2,51	2,80	1,75	2,40	2,00
x3(3)	59,74	65,64	59,17	60,49	61,26	61,03	61,11	61,80
x3(4)	1,12	1,08	1,06	1,01	1,03	0,88	0,97	0,89
x4(1)	8,37	8,24	8,13	7,41	7,49	7,60	7,39	6,91
x4(2)	26,83	24,37	30,55	30,40	30,24	31,62	32,87	37,94
x4(3)	2,53	2,87	2,93	2,90	2,99	2,87	3,49	3,39
x4(4)	1,12	1,06	0,82	0,45	0,44	0,61	0,49	0,82
x4(5)	71,37	69,91	71,43	70,41	70,33	68,50	68,41	66,10
x4(6)	16,28	12,75	16,07	12,94	15,06	13,06	15,51	14,80
x4(7)	4,72	3,98	5,00	2,89	2,79	4,23	4,68	4,25
x5(1)	57,58	64,67	68,63	60,82	61,92	58,43	66,32	64,67
x5(2)	85,68	88,40	86,05	88,49	87,86	85,54	85,64	88,40
x5(3)	0,93	1,00	1,05	1,10	1,10	0,91	1,05	1,00
x5(4)	95,14	90,89	92,27	94,01	92,84	93,97	92,69	90,89
x5(5)	14,70	14,02	14,80	14,21	16,79	14,69	14,80	14,02
x5(6)	14,00	15,01	13,80	14,20	15,20	14,00	13,80	15,01
x5(7)	6,10	5,54	6,13	5,39	5,88	5,76	6,06	5,54
x6(1)	66,25	63,34	66,01	55,97	55,51	59,03	71,37	74,68
x6(2)	29,65	28,48	27,45	30,18	29,53	30,49	30,12	30,72
x6(3)	4,17	6,71	6,44	18,71	6,98	3,96	12,86	4,59
x6(4)	15,78	17,65	17,81	17,44	16,88	15,92	17,88	17,85
x6(5)	47,96	47,83	46,05	44,18	52,69	52,79	37,88	38,43
x6(6)	1,19	1,18	1,20	0,98	1,03	1,09	0,98	0,97
x6(7)	92,88	92,33	93,21	91,57	90,81	91,82	94,78	94,72
x7(1)	5,86	7,35	4,45	7,70	8,71	6,25	5,56	7,24
x7(2)	2,74	2,55	2,46	2,95	2,10	2,81	2,97	3,00
x7(3)	1,35	1,46	0,92	1,61	1,20	1,57	1,19	1,57
x7(4)	3,61	1,38	1,27	4,76	2,34	4,51	3,38	6,09
x7(5)	2,50	2,45	2,21	2,82	2,04	2,69	2,89	2,89
x7(6)	1,93	2,50	1,30	2,16	2,92	2,49	2,28	2,33
x8(1)	5,87	4,39	4,00	5,91	5,53	6,86	6,12	5,44
x8(2)	82,81	78,26	82,63	81,78	77,95	85,15	82,32	82,27

Tabela C.1. (cont.) – Valores das Variáveis de Entrada para Simulação

Variáveis	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Caso 6	Caso 7	Caso 8
x8(3)	0,85	0,79	0,90	1,80	1,40	0,95	1,00	0,90
x8(4)	1,68	1,63	1,26	1,49	1,59	1,65	1,66	1,58
x9(1)	2,30	2,67	2,53	2,52	2,37	2,28	1,47	2,44
x9(2)	29,12	29,05	30,11	28,57	28,80	29,22	29,23	30,03
x9(3)	11,56	13,05	13,02	12,68	11,89	11,69	9,53	11,55
x9(4)	45,94	53,92	53,39	50,98	46,97	45,27	38,43	46,99
x9(5)	102,12	102,28	103,24	103,65	103,58	103,68	102,52	102,31
x9(6)	68,13	70,68	72,55	70,85	69,80	69,24	66,23	69,12
X10(1)	61,32	61,58	61,40	63,77	61,16	60,91	61,05	61,81
X10(2)	2,02	2,34	2,20	2,60	2,29	2,31	2,30	2,69
X10(3)	13,01	13,22	13,60	14,60	12,52	13,69	14,20	15,00
X10(4)	12,04	13,05	13,47	12,47	18,19	21,31	22,62	17,56
X10(5)	14,34	18,16	10,34	20,53	10,46	10,23	8,99	17,82
X10(6)	62,28	64,19	59,30	59,37	57,54	58,08	57,96	57,75

Tabela C.2. – Resultados da Simulação do Processo Global**Caso 1**

x1(1)	14,02	y1(1)	14,30	x6(1)	66,25	y6(1)	1,77
x1(2)	14,81	y1(2)	1,33	x6(2)	29,65	y6(2)	1,46
x1(3)	4,35			x6(3)	4,17		
x1(4)	71,17			x6(4)	15,78		
				x6(5)	47,96		
x2(1)	14,02	y2(1)	7,94	x6(6)	1,19		
x2(2)	65,63	y2(2)	24,79	x6(7)	92,88		
x2(3)	14,30	y2(3)	2,74				
x2(4)	3,02	y2(4)	0,93	x7(1)	8,11	y7(1)	2,25
x2(5)	50,15	y2(5)	5,35	x7(2)	2,50	y7(2)	27,52
x2(6)	11,12			x7(3)	1,21	y7(3)	5,57
				x7(4)	3,47		
x3(1)	5,35	y3(1)	3,14	x7(5)	2,50		
x3(2)	2,80	y3(2)	49,29	x7(6)	1,93		
x3(3)	59,74	y3(3)	2,20				
x3(4)	1,12	y3(4)	119,83	x8(1)	5,57	y8(1)	1,68
				x8(2)	82,81	y8(2)	106,44
x4(1)	7,94	y4(1)	3,81	x8(3)	0,85	y8(3)	29,18
x4(2)	24,79	y4(2)	3,52	x8(4)	1,68	y8(4)	109,85
x4(3)	2,74	y4(3)	25,81				
x4(4)	0,93	y4(4)	39,39	x9(1)	2,25	y9(1)	40,08
x4(5)	57,10	y4(5)	54,82	x9(2)	27,52	y9(2)	59,05
x4(6)	16,28	y4(6)	88,19	x9(3)	11,56	y9(3)	1,94
x4(7)	4,72			x9(4)	45,94		
				x9(5)	102,12		
x5(1)	54,82	y5(1)	2,50	x9(6)	68,13		
x5(2)	88,19	y5(2)	1,21				
x5(3)	1,19	y5(3)	3,47	x10(1)	59,05	y10(1)	1,77
x5(4)	92,88	y5(4)	1,75	x10(2)	1,94	y10(2)	0,18
x5(5)	14,70	y5(5)	99,99	x10(3)	13,01	y10(3)	52,81
x5(6)	14,00	y5(6)	69,92	x10(4)	12,04		
x5(7)	6,10	y5(7)	103,57	x10(5)	14,34		
		y5(8)	17,28	x10(6)	62,28		

Número de Iterações: 7

Tabela C.2 (cont.) – Resultados da Simulação do Processo Global**Caso 2**

x1(1)	13,99	y1(1)	20,09	x6(1)	63,34	y6(1)	1,77
x1(2)	20,72	y1(2)	8,17	x6(2)	28,48	y6(2)	1,50
x1(3)	14,90			x6(3)	6,71		
x1(4)	59,80			x6(4)	17,65		
				x6(5)	47,83		
x2(1)	13,99	y2(1)	7,88	x6(6)	1,18		
x2(2)	62,61	y2(2)	30,63	x6(7)	92,33		
x2(3)	20,09	y2(3)	3,34				
x2(4)	6,73	y2(4)	0,59	x7(1)	7,82	y7(1)	2,27
x2(5)	44,24	y2(5)	5,63	x7(2)	2,51	y7(2)	29,23
x2(6)	9,40			x7(3)	1,43	y7(3)	6,15
				x7(4)	4,12		
x3(1)	5,63	y3(1)	2,81	x7(5)	2,45		
x3(2)	2,80	y3(2)	51,11	x7(6)	2,50		
x3(3)	65,64	y3(3)	2,75				
x3(4)	1,08	y3(4)	125,96	x8(1)	6,15	y8(1)	1,67
				x8(2)	78,26	y8(2)	100,63
x4(1)	7,88	y4(1)	3,53	x8(3)	0,79	y8(3)	20,30
x4(2)	30,63	y4(2)	3,49	x8(4)	1,63	y8(4)	105,45
x4(3)	3,34	y4(3)	29,77				
x4(4)	0,59	y4(4)	38,81	x9(1)	2,27	y9(1)	39,98
x4(5)	56,10	y4(5)	53,53	x9(2)	29,23	y9(2)	62,09
x4(6)	12,75	y4(6)	85,96	x9(3)	13,05	y9(3)	1,90
x4(7)	3,98			x9(4)	53,92		
				x9(5)	102,28		
x5(1)	53,53	y5(1)	2,51	x9(6)	70,68		
x5(2)	85,96	y5(2)	1,43				
x5(3)	1,18	y5(3)	4,12	x10(1)	62,09	y10(1)	1,87
x5(4)	92,33	y5(4)	1,72	x10(2)	1,90	y10(2)	0,37
x5(5)	14,02	y5(5)	98,70	x10(3)	13,22	y10(3)	52,61
x5(6)	15,01	y5(6)	68,25	x10(4)	13,05		
x5(7)	5,54	y5(7)	102,60	x10(5)	18,16		
		y5(8)	14,71	x10(6)	64,19		

Número de Iterações: 6

Tabela C.2 (cont.) – Resultados da Simulação do Processo Global

Caso 3							
x1(1)	14,50	y1(1)	17,23	x6(1)	66,01	y6(1)	1,75
x1(2)	18,38	y1(2)	1,46	x6(2)	27,45	y6(2)	1,50
x1(3)	5,24			x6(3)	6,44		
x1(4)	68,16			x6(4)	17,81		
				x6(5)	46,05		
x2(1)	14,50	y2(1)	8,38	x6(6)	1,20		
x2(2)	65,19	y2(2)	28,85	x6(7)	93,21		
x2(3)	17,23	y2(3)	2,60				
x2(4)	3,78	y2(4)	0,97	x7(1)	7,67	y7(1)	2,26
x2(5)	46,18	y2(5)	5,93	x7(2)	2,59	y7(2)	29,09
x2(6)	10,01			x7(3)	1,32	y7(3)	4,07
				x7(4)	2,53		
x3(1)	5,93	y3(1)	3,74	x7(5)	2,21		
x3(2)	2,30	y3(2)	49,62	x7(6)	1,30		
x3(3)	59,17	y3(3)	2,04				
x3(4)	1,06	y3(4)	125,00	x8(1)	4,07	y8(1)	1,18
				x8(2)	82,63	y8(2)	103,49
x4(1)	8,38	y4(1)	3,72	x8(3)	0,90	y8(3)	23,71
x4(2)	28,85	y4(2)	3,20	x8(4)	1,26	y8(4)	108,53
x4(3)	2,60	y4(3)	26,58				
x4(4)	0,97	y4(4)	39,26	x9(1)	2,26	y9(1)	41,96
x4(5)	57,82	y4(5)	56,65	x9(2)	29,09	y9(2)	61,94
x4(6)	16,07	y4(6)	88,42	x9(3)	13,02	y9(3)	1,93
x4(7)	5,00			x9(4)	53,39		
				x9(5)	103,24		
x5(1)	56,65	y5(1)	2,59	x9(6)	72,55		
x5(2)	88,42	y5(2)	1,32				
x5(3)	1,20	y5(3)	2,53	x10(1)	61,94	y10(1)	1,80
x5(4)	93,21	y5(4)	1,78	x10(2)	1,93	y10(2)	0,26
x5(5)	14,80	y5(5)	100,36	x10(3)	13,60	y10(3)	52,79
x5(6)	13,80	y5(6)	71,25	x10(4)	13,47		
x5(7)	6,13	y5(7)	103,82	x10(5)	10,34		
		y5(8)	17,90	x10(6)	59,30		

Número de Iterações: 10

Tabela C.2 (cont.) – Resultados da Simulação do Processo Global

Caso 4

x1(1)	12,49	y1(1)	19,61	x6(1)	55,97	y6(1)	1,54
x1(2)	20,20	y1(2)	3,76	x6(2)	30,18	y6(2)	1,41
x1(3)	4,63			x6(3)	18,71		
x1(4)	70,06			x6(4)	17,44		
				x6(5)	44,18		
x2(1)	12,49	y2(1)	6,26	x6(6)	0,98		
x2(2)	65,70	y2(2)	37,72	x6(7)	91,57		
x2(3)	19,61	y2(3)	1,08				
x2(4)	0,87	y2(4)	0,16	x7(1)	6,71	y7(1)	2,39
x2(5)	48,12	y2(5)	6,31	x7(2)	2,63	y7(2)	28,89
x2(6)	10,01			x7(3)	1,26	y7(3)	6,12
				x7(4)	4,27		
x3(1)	6,31	y3(1)	3,19	x7(5)	2,82		
x3(2)	2,51	y3(2)	49,95	x7(6)	2,16		
x3(3)	60,49	y3(3)	2,99				
x3(4)	1,01	y3(4)	123,64	x8(1)	6,12	y8(1)	1,65
				x8(2)	81,78	y8(2)	103,57
x4(1)	6,26	y4(1)	2,54	x8(3)	1,80	y8(3)	22,22
x4(2)	37,72	y4(2)	3,41	x8(4)	1,49	y8(4)	106,31
x4(3)	1,08	y4(3)	27,90				
x4(4)	0,16	y4(4)	34,64	x9(1)	2,39	y9(1)	41,36
x4(5)	58,60	y4(5)	59,61	x9(2)	28,89	y9(2)	62,05
x4(6)	12,94	y4(6)	86,46	x9(3)	12,68	y9(3)	2,08
x4(7)	2,89			x9(4)	50,98		
				x9(5)	103,65		
x5(1)	59,61	y5(1)	2,63	x9(6)	70,85		
x5(2)	86,46	y5(2)	1,26				
x5(3)	0,98	y5(3)	4,27	x10(1)	62,05	y10(1)	1,94
x5(4)	91,57	y5(4)	1,59	x10(2)	2,08	y10(2)	0,08
x5(5)	14,21	y5(5)	98,72	x10(3)	14,60	y10(3)	53,45
x5(6)	14,20	y5(6)	70,96	x10(4)	12,47		
x5(7)	5,39	y5(7)	102,35	x10(5)	20,53		
		y5(8)	19,21	x10(6)	59,37		

Número de Iterações: 16

Tabela C.2 (cont.) – Resultados da Simulação do Processo Global**Caso 5**

x1(1)	12,50	y1(1)	19,67	x6(1)	55,51	y6(1)	1,59
x1(2)	20,60	y1(2)	8,90	x6(2)	29,53	y6(2)	1,49
x1(3)	9,20			x6(3)	6,98		
x1(4)	73,05			x6(4)	16,88		
				x6(5)	52,69		
x2(1)	12,50	y2(1)	6,27	x6(6)	1,03		
x2(2)	64,80	y2(2)	37,15	x6(7)	90,81		
x2(3)	19,67	y2(3)	0,80				
x2(4)	0,30	y2(4)	0,45	x7(1)	6,79	y7(1)	2,47
x2(5)	44,53	y2(5)	6,38	x7(2)	2,76	y7(2)	27,38
x2(6)	9,16			x7(3)	1,29	y7(3)	6,83
				x7(4)	6,08		
x3(1)	6,38	y3(1)	2,88	x7(5)	2,04		
x3(2)	2,80	y3(2)	49,79	x7(6)	2,92		
x3(3)	61,26	y3(3)	3,08				
x3(4)	1,03	y3(4)	122,75	x8(1)	6,83	y8(1)	1,79
				x8(2)	77,95	y8(2)	100,63
x4(1)	6,27	y4(1)	2,16	x8(3)	1,40	y8(3)	20,24
x4(2)	37,15	y4(2)	3,86	x8(4)	1,59	y8(4)	105,54
x4(3)	0,80	y4(3)	28,77				
x4(4)	0,45	y4(4)	34,36	x9(1)	2,47	y9(1)	40,98
x4(5)	54,65	y4(5)	53,76	x9(2)	27,38	y9(2)	60,13
x4(6)	15,06	y4(6)	85,92	x9(3)	11,89	y9(3)	2,15
x4(7)	2,79			x9(4)	46,97		
				x9(5)	103,58		
x5(1)	53,76	y5(1)	2,76	x9(6)	69,80		
x5(2)	85,92	y5(2)	1,29				
x5(3)	1,03	y5(3)	6,08	x10(1)	60,13	y10(1)	1,98
x5(4)	90,81	y5(4)	1,64	x10(2)	2,15	y10(2)	0,14
x5(5)	16,79	y5(5)	98,20	x10(3)	12,52	y10(3)	51,04
x5(6)	15,20	y5(6)	69,45	x10(4)	18,19		
x5(7)	5,88	y5(7)	102,19	x10(5)	10,46		
		y5(8)	15,28	x10(6)	57,54		

Número de Iterações:12

Tabela C.2 (cont.) – Resultados da Simulação do Processo Global

Caso 6

x1(1)	9,50	y1(1)	17,21	x6(1)	59,03	y6(1)	1,67
x1(2)	18,72	y1(2)	17,17	x6(2)	30,49	y6(2)	1,51
x1(3)	18,77			x6(3)	3,96		
x1(4)	65,97			x6(4)	15,92		
				x6(5)	52,79		
x2(1)	9,50	y2(1)	5,36	x6(6)	1,09		
x2(2)	63,59	y2(2)	35,88	x6(7)	91,82		
x2(3)	17,21	y2(3)	0,73				
x2(4)	1,60	y2(4)	0,71	x7(1)	5,87	y7(1)	2,20
x2(5)	39,07	y2(5)	4,43	x7(2)	2,44	y7(2)	26,41
x2(6)	8,99			x7(3)	1,13	y7(3)	7,07
				x7(4)	6,16		
x3(1)	4,43	y3(1)	3,68	x7(5)	2,69		
x3(2)	1,75	y3(2)	51,70	x7(6)	2,49		
x3(3)	61,03	y3(3)	1,21				
x3(4)	0,88	y3(4)	128,85	x8(1)	7,07	y8(1)	1,62
				x8(2)	85,15	y8(2)	104,29
x4(1)	5,36	y4(1)	1,67	x8(3)	0,95	y8(3)	26,82
x4(2)	35,88	y4(2)	3,42	x8(4)	1,65	y8(4)	108,63
x4(3)	0,73	y4(3)	39,33				
x4(4)	0,71	y4(4)	35,64	x9(1)	2,20	y9(1)	40,09
x4(5)	64,12	y4(5)	65,51	x9(2)	26,41	y9(2)	59,15
x4(6)	13,06	y4(6)	91,78	x9(3)	11,69	y9(3)	1,80
x4(7)	4,23			x9(4)	45,27		
				x9(5)	103,68		
x5(1)	65,50	y5(1)	2,44	x9(6)	69,24		
x5(2)	91,77	y5(2)	1,13				
x5(3)	1,09	y5(3)	6,16	x10(1)	59,15	y10(1)	1,77
x5(4)	91,82	y5(4)	1,62	x10(2)	1,80	y10(2)	0,07
x5(5)	14,69	y5(5)	99,06	x10(3)	13,69	y10(3)	53,05
x5(6)	14,00	y5(6)	76,99	x10(4)	21,31		
x5(7)	5,76	y5(7)	103,31	x10(5)	10,23		
		y5(8)	21,99	x10(6)	58,08		

Número de Iterações: 16

Tabela C.2 (cont.) – Resultados da Simulação do Processo Global**Caso 7**

x1(1)	11,96	y1(1)	18,47	x6(1)	71,37	y6(1)	1,52
x1(2)	19,34	y1(2)	16,26	x6(2)	30,12	y6(2)	1,42
x1(3)	18,66			x6(3)	12,86		
x1(4)	60,09			x6(4)	17,88		
				x6(5)	37,88		
x2(1)	11,96	y2(1)	7,37	x6(6)	0,98		
x2(2)	63,17	y2(2)	30,17	x6(7)	94,78		
x2(3)	18,47	y2(3)	2,37				
x2(4)	2,40	y2(4)	0,64	x7(1)	7,53	y7(1)	2,48
x2(5)	34,80	y2(5)	4,61	x7(2)	2,77	y7(2)	29,61
x2(6)	8,65			x7(3)	1,43	y7(3)	6,11
				x7(4)	0,86		
x3(1)	4,61	y3(1)	2,85	x7(5)	2,89		
x3(2)	2,40	y3(2)	49,85	x7(6)	2,28		
x3(3)	61,11	y3(3)	1,79				
x3(4)	0,97	y3(4)	120,53	x8(1)	6,11	y8(1)	1,68
				x8(2)	82,32	y8(2)	104,12
x4(1)	7,37	y4(1)	2,85	x8(3)	1,00	y8(3)	24,76
x4(2)	30,17	y4(2)	3,94	x8(4)	1,66	y8(4)	108,15
x4(3)	2,37	y4(3)	30,98				
x4(4)	0,64	y4(4)	38,23	x9(1)	2,48	y9(1)	33,85
x4(5)	66,14	y4(5)	67,92	x9(2)	29,61	y9(2)	60,78
x4(6)	15,51	y4(6)	90,02	x9(3)	9,53	y9(3)	2,68
x4(7)	4,68			x9(4)	38,43		
				x9(5)	102,52		
x5(1)	67,92	y5(1)	2,77	x9(6)	66,23		
x5(2)	90,02	y5(2)	1,43				
x5(3)	0,98	y5(3)	0,86	x10(1)	60,78	y10(1)	2,39
x5(4)	94,78	y5(4)	1,51	x10(2)	2,68	y10(2)	0,37
x5(5)	14,80	y5(5)	100,88	x10(3)	14,20	y10(3)	52,54
x5(6)	13,80	y5(6)	79,06	x10(4)	22,62		
x5(7)	6,06	y5(7)	104,18	x10(5)	8,99		
		y5(8)	19,94	x10(6)	57,96		

Número de Iterações: 7

Tabela C.2 (cont.) – Resultados da Simulação do Processo Global

Caso 8

x1(1)	11,96	y1(1)	18,47	x6(1)	74,68	y6(1)	1,47
x1(2)	19,34	y1(2)	16,26	x6(2)	30,72	y6(2)	1,32
x1(3)	18,66			x6(3)	4,59		
x1(4)	60,09			x6(4)	17,85		
				x6(5)	38,43		
x2(1)	11,96	y2(1)	7,39	x6(6)	0,97		
x2(2)	63,68	y2(2)	30,56	x6(7)	94,72		
x2(3)	18,47	y2(3)	2,18				
x2(4)	2,40	y2(4)	0,60	x7(1)	7,44	y7(1)	2,44
x2(5)	36,36	y2(5)	4,63	x7(2)	2,71	y7(2)	28,78
x2(6)	8,82			x7(3)	1,24	y7(3)	6,25
				x7(4)	1,84		
x3(1)	4,63	y3(1)	3,38	x7(5)	2,89		
x3(2)	2,00	y3(2)	51,48	x7(6)	2,33		
x3(3)	61,80	y3(3)	1,36				
x3(4)	0,89	y3(4)	126,24	x8(1)	6,25	y8(1)	1,62
				x8(2)	82,27	y8(2)	102,66
x4(1)	7,39	y4(1)	2,67	x8(3)	0,90	y8(3)	23,91
x4(2)	30,56	y4(2)	4,07	x8(4)	1,58	y8(4)	107,44
x4(3)	2,18	y4(3)	30,01				
x4(4)	0,60	y4(4)	37,60	x9(1)	2,44	y9(1)	41,94
x4(5)	66,62	y4(5)	68,07	x9(2)	28,78	y9(2)	59,73
x4(6)	14,80	y4(6)	88,94	x9(3)	11,55	y9(3)	2,17
x4(7)	4,25			x9(4)	46,99		
				x9(5)	102,31		
x5(1)	68,07	y5(1)	2,71	x9(6)	69,12		
x5(2)	88,94	y5(2)	1,24				
x5(3)	0,97	y5(3)	1,84	x10(1)	59,73	y10(1)	2,07
x5(4)	94,72	y5(4)	1,43	x10(2)	2,17	y10(2)	0,34
x5(5)	14,02	y5(5)	99,85	x10(3)	15,00	y10(3)	49,30
x5(6)	15,01	y5(6)	78,81	x10(4)	17,56		
x5(7)	5,54	y5(7)	103,23	x10(5)	17,82		
		y5(8)	24,90	x10(6)	57,75		

Número de Iterações: 8

ANEXO D – ANÁLISE DE SENSIBILIDADE

TABELAS DE RESULTADOS

As tabelas a seguir apresentam os resultados obtidos no estudo de análise de sensibilidade das variáveis do reator (REDE-1). As tabelas correspondem aos gráficos das Figuras 4.39 a 4.46 apresentadas no texto no item 4.5.5. Os resultados mostram a influência da temperatura e do teor de CPD da carga do reator na conversão do CPD e na perda de IP para diferentes especificações de vazão de carga e teor de IP na carga. As unidades das variáveis são arbitrárias.

Tabela D.1. – Teor de CPD no Reator – y1(2) – Caso 1

i \ j	56	58	60	62	64	66	68	70	72	74	76	78	80
2,5	2,67	2,47	2,33	2,25	2,20	2,18	2,16	2,15	2,13	2,12	2,10	2,06	2,01
3,5	3,29	3,05	2,88	2,78	2,73	2,69	2,68	2,66	2,65	2,64	2,61	2,58	2,51
4,5	3,82	3,53	3,33	3,21	3,15	3,11	3,09	3,08	3,07	3,05	3,03	2,98	2,90
5,5	4,25	3,91	3,68	3,54	3,46	3,42	3,40	3,39	3,38	3,36	3,34	3,28	3,19
6,5	4,58	4,20	3,93	3,78	3,69	3,65	3,63	3,62	3,61	3,59	3,55	3,49	3,39
7,5	4,81	4,41	4,12	3,95	3,86	3,81	3,79	3,78	3,77	3,75	3,71	3,64	3,52
8,5	4,95	4,55	4,25	4,08	3,99	3,95	3,93	3,92	3,91	3,89	3,85	3,76	3,62
9,5	5,01	4,64	4,37	4,21	4,13	4,10	4,08	4,08	4,07	4,04	3,99	3,90	3,74
10,5	4,99	4,70	4,48	4,35	4,30	4,28	4,28	4,28	4,27	4,24	4,19	4,08	3,91
11,5	4,90	4,73	4,59	4,53	4,51	4,52	4,53	4,54	4,54	4,51	4,45	4,34	4,15
12,5	4,77	4,74	4,71	4,72	4,77	4,82	4,86	4,88	4,88	4,86	4,80	4,68	4,47
13,5	4,60	4,71	4,82	4,94	5,07	5,17	5,25	5,29	5,31	5,29	5,23	5,10	4,88
14,5	4,43	4,66	4,91	5,16	5,39	5,57	5,70	5,77	5,81	5,80	5,73	5,60	5,37
15,5	4,33	4,62	4,98	5,37	5,72	6,00	6,19	6,31	6,37	6,37	6,31	6,18	5,94
16,5	4,32	4,61	5,04	5,55	6,04	6,43	6,70	6,88	6,97	7,00	6,95	6,81	6,57
17,5	4,45	4,67	5,11	5,71	6,32	6,84	7,23	7,48	7,62	7,67	7,63	7,50	7,25
18,5	4,73	4,84	5,23	5,86	6,58	7,23	7,73	8,07	8,28	8,36	8,35	8,22	7,97
19,5	5,16	5,15	5,43	6,02	6,80	7,57	8,20	8,66	8,94	9,08	9,09	8,98	8,73

Notas: i – % CPD na carga, x1(3)
 j – Temperatura da carga, x1(4)
 Vazão de carga, x1(1) = 14,02
 % IP na carga, x1(2) = 14,81

Tabela D.2. – Perda de IP no Reator – x1(2) - y1(1) – Caso 1

i \ j	56	58	60	62	64	66	68	70	72	74	76	78	80
2,5	1,20	1,33	1,32	1,23	1,09	0,93	0,78	0,66	0,59	0,62	0,81	1,25	2,04
3,5	1,02	1,18	1,19	1,11	0,98	0,83	0,69	0,58	0,54	0,59	0,80	1,28	2,12
4,5	0,87	1,04	1,08	1,01	0,89	0,76	0,63	0,53	0,50	0,57	0,82	1,33	2,23
5,5	0,75	0,93	0,98	0,93	0,82	0,70	0,58	0,50	0,48	0,57	0,84	1,40	2,34
6,5	0,66	0,84	0,90	0,86	0,77	0,65	0,54	0,47	0,47	0,58	0,89	1,48	2,48
7,5	0,60	0,78	0,84	0,81	0,72	0,61	0,51	0,45	0,47	0,60	0,94	1,58	2,62
8,5	0,59	0,74	0,80	0,76	0,68	0,58	0,49	0,44	0,47	0,63	1,00	1,68	2,77
9,5	0,62	0,74	0,77	0,74	0,65	0,56	0,48	0,44	0,48	0,66	1,06	1,79	2,92
10,5	0,72	0,77	0,78	0,73	0,64	0,54	0,46	0,43	0,49	0,70	1,13	1,90	3,08
11,5	0,87	0,85	0,82	0,74	0,64	0,53	0,46	0,43	0,51	0,73	1,20	2,02	3,23
12,5	1,07	0,99	0,90	0,78	0,65	0,54	0,46	0,44	0,53	0,78	1,28	2,14	3,39
13,5	1,34	1,18	1,03	0,86	0,69	0,55	0,46	0,45	0,55	0,82	1,36	2,26	3,55
14,5	1,64	1,43	1,21	0,98	0,76	0,59	0,49	0,47	0,57	0,87	1,44	2,39	3,71
15,5	1,95	1,72	1,45	1,15	0,88	0,66	0,53	0,50	0,61	0,92	1,53	2,53	3,87
16,5	2,26	2,04	1,73	1,38	1,03	0,76	0,59	0,54	0,65	0,98	1,62	2,66	4,03
17,5	2,54	2,36	2,06	1,66	1,25	0,91	0,69	0,61	0,71	1,06	1,73	2,81	4,19
18,5	2,77	2,67	2,40	1,98	1,52	1,10	0,82	0,70	0,79	1,14	1,84	2,96	4,34
19,5	2,96	2,94	2,73	2,34	1,84	1,35	1,00	0,83	0,89	1,24	1,97	3,11	4,50

Notas: i – % CPD na carga, x1(3)
 j – Temperatura da carga, x1(4)
 Vazão de carga, x1(1) = 14,02
 % IP na carga, x1(2) = 14,81

Tabela D.3. – Teor de CPD no Reator – y1(2) – Caso 2

i \ j	56	58	60	62	64	66	68	70	72	74	76	78	80
2,5	2,71	2,62	2,55	2,50	2,44	2,37	2,29	2,17	1,99	1,74	1,37	0,89	0,34
3,5	3,25	3,15	3,07	3,01	2,94	2,86	2,76	2,61	2,40	2,08	1,64	1,07	0,43
4,5	3,67	3,56	3,48	3,41	3,33	3,24	3,12	2,94	2,68	2,30	1,77	1,11	0,39
5,5	3,97	3,86	3,78	3,71	3,62	3,52	3,37	3,16	2,85	2,40	1,79	1,04	0,24
6,5	4,15	4,06	3,98	3,91	3,83	3,72	3,55	3,30	2,93	2,41	1,72	0,90	0,05
7,5	4,21	4,16	4,11	4,06	3,98	3,86	3,67	3,39	2,97	2,39	1,62	0,74	-0,14
8,5	4,17	4,19	4,19	4,17	4,10	3,98	3,78	3,47	3,01	2,37	1,55	0,64	-0,25
9,5	4,03	4,15	4,24	4,27	4,23	4,12	3,91	3,58	3,08	2,39	1,54	0,61	-0,27
10,5	3,81	4,08	4,27	4,38	4,39	4,30	4,09	3,74	3,22	2,50	1,63	0,70	-0,16
11,5	3,53	3,96	4,29	4,50	4,58	4,52	4,33	3,97	3,43	2,69	1,81	0,90	0,08
12,5	3,23	3,80	4,29	4,63	4,80	4,80	4,63	4,28	3,73	2,98	2,11	1,22	0,44
13,5	2,95	3,63	4,27	4,76	5,04	5,11	4,99	4,65	4,10	3,36	2,50	1,64	0,90
14,5	2,73	3,45	4,23	4,88	5,29	5,46	5,39	5,08	4,55	3,81	2,97	2,14	1,45
15,5	2,63	3,32	4,18	4,97	5,53	5,81	5,82	5,56	5,05	4,33	3,51	2,72	2,07
16,5	2,69	3,28	4,14	5,03	5,74	6,16	6,26	6,06	5,59	4,90	4,11	3,36	2,75
17,5	2,92	3,36	4,15	5,08	5,91	6,48	6,70	6,58	6,16	5,51	4,76	4,05	3,47
18,5	3,32	3,58	4,24	5,15	6,06	6,76	7,11	7,09	6,74	6,14	5,44	4,77	4,24
19,5	3,86	3,96	4,45	5,26	6,19	7,00	7,48	7,58	7,32	6,78	6,14	5,52	5,03

Notas: i – % CPD na carga, x1(3)
 j – Temperatura da carga, x1(4)
 Vazão de carga, x1(1) = 14,02
 % IP na carga, x1(2) = 12,00

Tabela D.4. – Perda de IP no Reator – x1(2) - y1(1) – Caso 2

i \ j	56	58	60	62	64	66	68	70	72	74	76	78	80
2,5	2,58	2,52	2,46	2,43	2,44	2,54	2,74	3,01	3,27	3,43	3,49	3,50	3,50
3,5	2,47	2,41	2,35	2,32	2,35	2,47	2,70	3,00	3,28	3,44	3,49	3,50	3,50
4,5	2,36	2,30	2,24	2,22	2,27	2,41	2,67	3,00	3,28	3,44	3,49	3,50	3,50
5,5	2,25	2,19	2,13	2,12	2,18	2,35	2,65	3,00	3,29	3,45	3,49	3,50	3,50
6,5	2,15	2,08	2,03	2,02	2,10	2,30	2,63	3,01	3,31	3,45	3,50	3,50	3,50
7,5	2,07	1,98	1,92	1,93	2,02	2,25	2,61	3,02	3,32	3,46	3,50	3,50	3,50
8,5	2,00	1,90	1,83	1,84	1,95	2,21	2,61	3,03	3,33	3,46	3,50	3,50	3,50
9,5	1,96	1,83	1,75	1,76	1,89	2,18	2,60	3,04	3,34	3,47	3,50	3,50	3,50
10,5	1,94	1,79	1,69	1,69	1,84	2,15	2,61	3,06	3,36	3,47	3,50	3,50	3,50
11,5	1,95	1,78	1,65	1,64	1,79	2,13	2,61	3,08	3,37	3,48	3,50	3,50	3,50
12,5	2,00	1,80	1,65	1,62	1,77	2,13	2,63	3,10	3,38	3,48	3,50	3,50	3,50
13,5	2,06	1,85	1,67	1,62	1,77	2,14	2,65	3,13	3,40	3,48	3,50	3,50	3,50
14,5	2,14	1,93	1,73	1,65	1,79	2,16	2,69	3,16	3,41	3,49	3,50	3,50	3,50
15,5	2,22	2,03	1,83	1,72	1,83	2,20	2,73	3,18	3,42	3,49	3,50	3,50	3,51
16,5	2,30	2,14	1,94	1,82	1,91	2,26	2,77	3,21	3,43	3,49	3,50	3,50	3,51
17,5	2,36	2,24	2,07	1,95	2,02	2,34	2,83	3,24	3,44	3,49	3,50	3,50	3,51
18,5	2,41	2,34	2,21	2,10	2,15	2,44	2,90	3,28	3,45	3,50	3,50	3,50	3,51
19,5	2,43	2,42	2,33	2,25	2,29	2,55	2,97	3,31	3,46	3,50	3,50	3,50	3,51

Notas: i – % CPD na carga, x1(3)
 j – Temperatura da carga, x1(4)
 Vazão de carga, x1(1) = 14,02
 % IP na carga, x1(2) = 12,00

Tabela D.5. - Teor de CPD no Reator – y1(2) – Caso 3

i \ j	56	58	60	62	64	66	68	70	72	74	76	78	80
2,5	3,15	2,87	2,62	2,44	2,33	2,27	2,24	2,23	2,23	2,23	2,24	2,24	2,24
3,5	3,92	3,59	3,29	3,07	2,94	2,87	2,83	2,82	2,82	2,83	2,83	2,84	2,85
4,5	4,63	4,25	3,89	3,62	3,46	3,37	3,33	3,32	3,32	3,32	3,33	3,35	3,36
5,5	5,26	4,82	4,40	4,08	3,88	3,77	3,72	3,71	3,71	3,72	3,74	3,75	3,77
6,5	5,81	5,32	4,83	4,45	4,21	4,08	4,02	4,00	4,01	4,02	4,04	4,06	4,08
7,5	6,27	5,75	5,19	4,75	4,46	4,31	4,24	4,22	4,22	4,24	4,26	4,28	4,30
8,5	6,64	6,10	5,49	4,99	4,67	4,49	4,41	4,38	4,39	4,41	4,43	4,46	4,48
9,5	6,91	6,39	5,76	5,22	4,86	4,66	4,57	4,54	4,54	4,56	4,59	4,62	4,64
10,5	7,08	6,62	6,00	5,44	5,06	4,84	4,74	4,71	4,72	4,74	4,77	4,80	4,82
11,5	7,15	6,81	6,23	5,68	5,29	5,07	4,97	4,94	4,95	4,97	5,00	5,03	5,06
12,5	7,11	6,94	6,46	5,94	5,57	5,36	5,26	5,24	5,25	5,27	5,31	5,34	5,36
13,5	6,97	7,00	6,66	6,23	5,89	5,70	5,62	5,61	5,63	5,66	5,69	5,73	5,75
14,5	6,76	7,01	6,84	6,52	6,25	6,10	6,05	6,05	6,08	6,12	6,16	6,20	6,22
15,5	6,52	6,95	6,97	6,79	6,62	6,54	6,53	6,56	6,61	6,66	6,71	6,74	6,77
16,5	6,31	6,86	7,06	7,04	6,99	6,99	7,04	7,11	7,19	7,26	7,31	7,36	7,39
17,5	6,19	6,78	7,10	7,24	7,33	7,44	7,57	7,70	7,81	7,90	7,98	8,03	8,07
18,5	6,20	6,75	7,14	7,41	7,64	7,87	8,10	8,30	8,46	8,59	8,68	8,75	8,80
19,5	6,37	6,82	7,21	7,55	7,89	8,26	8,60	8,90	9,12	9,29	9,42	9,50	9,56

Notas: i – % CPD na carga, x1(3)
 j – Temperatura da carga, x1(4)
 Vazão de carga, x1(1) = 14,02
 % IP na carga, x1(2) = 17,50

Tabela D.6. – Perda de IP no Reator – x1(2) - y1(1) – Caso 3

i \ j	60	62	64	66	68	70	72	74	76	78	80
2,5	0,46	0,67	0,76	0,76	0,71	0,65	0,57	0,50	0,43	0,38	0,35
3,5	0,40	0,64	0,74	0,75	0,72	0,66	0,59	0,52	0,45	0,40	0,38
4,5	0,36	0,62	0,74	0,76	0,73	0,68	0,61	0,54	0,48	0,44	0,41
5,5	0,32	0,61	0,74	0,78	0,76	0,71	0,65	0,58	0,53	0,48	0,46
6,5	0,30	0,60	0,76	0,81	0,80	0,75	0,69	0,63	0,58	0,54	0,52
7,5	0,29	0,61	0,78	0,84	0,84	0,80	0,74	0,68	0,63	0,59	0,58
8,5	0,29	0,62	0,81	0,88	0,89	0,85	0,80	0,74	0,69	0,65	0,65
9,5	0,31	0,64	0,84	0,92	0,93	0,90	0,85	0,79	0,75	0,72	0,71
10,5	0,34	0,67	0,88	0,97	0,98	0,95	0,90	0,85	0,80	0,77	0,78
11,5	0,39	0,71	0,91	1,01	1,03	1,00	0,95	0,90	0,86	0,83	0,84
12,5	0,48	0,77	0,97	1,06	1,07	1,05	1,00	0,95	0,91	0,89	0,90
13,5	0,60	0,85	1,03	1,11	1,12	1,09	1,05	1,00	0,96	0,94	0,95
14,5	0,77	0,96	1,11	1,18	1,18	1,15	1,10	1,04	1,01	0,99	1,01
15,5	0,99	1,12	1,22	1,26	1,25	1,20	1,14	1,09	1,05	1,04	1,06
16,5	1,28	1,32	1,37	1,37	1,33	1,26	1,20	1,14	1,10	1,08	1,11
17,5	1,62	1,59	1,56	1,51	1,43	1,34	1,26	1,19	1,14	1,13	1,16
18,5	2,01	1,91	1,81	1,69	1,56	1,44	1,33	1,25	1,19	1,18	1,22
19,5	2,43	2,28	2,12	1,93	1,73	1,55	1,42	1,31	1,25	1,23	1,27

Notas: i – % CPD na carga, x1(3)
 j – Temperatura da carga, x1(4)
 Vazão de carga, x1(1) = 14,02
 % IP na carga, x1(2) = 17,50

Tabela D.7. - Teor de CPD no Reator – y1(2) – Caso 4

	56	58	60	62	64	66	68	70	72	74	76	78	80
2,5	3,45	3,24	2,97	2,70	2,51	2,39	2,33	2,31	2,32	2,33	2,34	2,36	2,38
3,5	4,27	4,03	3,73	3,42	3,18	3,04	2,97	2,95	2,95	2,97	2,99	3,02	3,04
4,5	5,02	4,77	4,42	4,06	3,77	3,60	3,51	3,49	3,50	3,52	3,55	3,58	3,61
5,5	5,70	5,43	5,04	4,61	4,27	4,06	3,95	3,92	3,93	3,96	4,00	4,04	4,08
6,5	6,29	6,01	5,57	5,08	4,67	4,41	4,28	4,24	4,26	4,30	4,34	4,39	4,44
7,5	6,76	6,49	6,01	5,45	4,97	4,66	4,51	4,46	4,48	4,53	4,58	4,64	4,70
8,5	7,11	6,87	6,37	5,75	5,20	4,83	4,65	4,60	4,62	4,67	4,74	4,81	4,88
9,5	7,32	7,14	6,65	5,99	5,38	4,96	4,75	4,68	4,71	4,77	4,85	4,93	5,01
10,5	7,38	7,32	6,87	6,19	5,53	5,07	4,83	4,76	4,78	4,86	4,95	5,04	5,13
11,5	7,28	7,40	7,04	6,38	5,70	5,20	4,94	4,86	4,89	4,97	5,07	5,18	5,28
12,5	7,04	7,39	7,16	6,56	5,89	5,38	5,10	5,02	5,05	5,14	5,25	5,36	5,47
13,5	6,67	7,28	7,24	6,75	6,12	5,61	5,33	5,24	5,28	5,38	5,50	5,63	5,75
14,5	6,22	7,08	7,28	6,94	6,38	5,90	5,63	5,55	5,59	5,70	5,83	5,97	6,10
15,5	5,73	6,82	7,26	7,11	6,67	6,24	5,98	5,92	5,97	6,09	6,24	6,39	6,53
16,5	5,31	6,52	7,19	7,26	6,96	6,60	6,39	6,35	6,42	6,55	6,71	6,88	7,03
17,5	5,00	6,23	7,08	7,37	7,24	6,99	6,83	6,82	6,92	7,07	7,25	7,43	7,60
18,5	4,88	6,02	6,95	7,43	7,49	7,37	7,29	7,32	7,45	7,63	7,83	8,03	8,21
19,5	4,96	5,94	6,86	7,47	7,70	7,73	7,74	7,84	8,01	8,22	8,45	8,67	8,87

Notas: i – % CPD na carga, x1(3)
 j – Temperatura da carga, x1(4)
 Vazão de carga, x1(1) = 14,02
 % IP na carga, x1(2) = 20,00

Tabela D.8. – Perda de IP no Reator – x1(2) - y1(1) – Caso 4

i \ j	58	60	62	64	66	68	70	72	74	76	78	80
2,5	0,33	0,78	1,21	1,54	1,75	1,87	1,93	1,95	1,95	1,94	1,93	1,92
3,5	0,27	0,71	1,15	1,50	1,73	1,87	1,93	1,96	1,97	1,97	1,96	1,95
4,5	0,21	0,65	1,10	1,47	1,72	1,87	1,95	1,99	2,00	2,00	1,99	1,98
5,5	0,17	0,59	1,05	1,44	1,71	1,88	1,97	2,02	2,04	2,04	2,04	2,03
6,5	0,13	0,54	1,00	1,41	1,71	1,89	2,00	2,05	2,07	2,08	2,08	2,08
7,5	0,11	0,50	0,96	1,38	1,70	1,90	2,02	2,08	2,12	2,13	2,13	2,13
8,5	0,09	0,46	0,92	1,35	1,69	1,91	2,04	2,12	2,16	2,17	2,18	2,18
9,5	0,09	0,43	0,88	1,32	1,68	1,92	2,06	2,15	2,19	2,22	2,23	2,23
10,5	0,10	0,41	0,84	1,29	1,66	1,91	2,08	2,17	2,22	2,25	2,27	2,28
11,5	0,13	0,40	0,80	1,25	1,63	1,91	2,08	2,19	2,25	2,28	2,31	2,32
12,5	0,19	0,41	0,78	1,22	1,61	1,89	2,08	2,20	2,27	2,31	2,33	2,35
13,5	0,27	0,43	0,77	1,19	1,58	1,88	2,08	2,20	2,28	2,32	2,36	2,38
14,5	0,39	0,48	0,77	1,17	1,55	1,86	2,06	2,20	2,28	2,33	2,37	2,39
15,5	0,54	0,57	0,80	1,16	1,53	1,84	2,05	2,19	2,28	2,33	2,38	2,41
16,5	0,75	0,70	0,86	1,17	1,52	1,82	2,04	2,18	2,27	2,33	2,38	2,41
17,5	1,00	0,87	0,96	1,22	1,53	1,82	2,03	2,16	2,26	2,32	2,37	2,41
18,5	1,29	1,09	1,10	1,29	1,57	1,83	2,02	2,16	2,25	2,31	2,36	2,40
19,5	1,60	1,35	1,29	1,41	1,64	1,86	2,03	2,15	2,24	2,30	2,35	2,39

Notas: i – % CPD na carga, x1(3)
 j – Temperatura da carga, x1(4)
 Vazão de carga, x1(1) = 14,02
 % IP na carga, x1(2) = 20,00

ANEXO E – Resultados da Otimização do Processo

As tabelas a seguir apresentam os resultados da otimização para o caso considerado como exemplo neste estudo. A terceira coluna de cada tabela mostra o resultado da simulação para o caso base considerado, antes da otimização conforme já mostrado na Tabela C.2. (ANEXO 2).As demais colunas mostram os casos selecionados como ótimos para atingir o menor consumo de energia.

Os casos grifados em vermelho são os casos considerados para a otimização do bloco seguinte, isto é, da seção de fracionamento.

Tabela E.1 – Resultados da Otimização da Seção de Preparação de Carga

Variáveis		Simulação	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5
Notação	Descrição						
x1(1)	Vazão Carga	13,99	10,76	10,76	12,01	12,01	13,27
x1(2)	% IP carga	20,72	20,72	20,72	20,72	20,72	20,72
x1(3)	% CPD carga	14,9	14,9	14,9	14,9	14,9	14,9
x1(4)	Temp. Carga	59,8	61,9	61,9	68,0	68,0	74,1
y1(1)	% IP reator	20,09	19,71	19,71	20,07	20,07	19,21
y1(2)	Δ CPD	8,17	13,74	13,74	12,41	12,41	12,16
x2(1)	Vazão Carga	13,99	10,76	10,76	12,01	12,01	13,27
x2(2)	Temp. Carga	62,61	62,61	62,61	62,61	62,61	62,61
x2(3)	% IP carga	20,09	19,71	19,71	20,07	20,07	19,21
x2(4)	% CPD carga	6,73	1,16	1,16	2,49	2,49	2,74
x2(5)	Refluxo	44,24	31,22	36,72	31,22	36,72	31,22
x2(6)	Vapor Referv.	9,4	7,9	7,9	7,9	7,9	7,9
y2(1)	Vazão Topo	7,88	6,29	5,63	7,49	7,17	7,46
y2(2)	% IP topo	30,63	30,78	33,23	30,83	29,71	32,65
y2(3)	% CPD topo	3,34	1,36	0,89	2,00	2,07	1,92
y2(4)	% 2M2B topo	0,59	0,79	0,86	0,96	0,88	1,10
y2(5)	Vazão Fundo	5,63	4,49	4,94	4,60	4,77	5,13
x3(1)	Vazão Carga	5,63	4,49	4,94	4,60	4,77	5,13
x3(2)	Refluxo	2,8	2,5	1,6	2,1	1,6	1,6
x3(3)	Temp. Carga	65,64	65,64	65,64	65,64	65,64	65,64
x3(4)	Vapor Referv.	1,08	0,83	0,83	0,83	0,83	0,83
y3(1)	Vazão Topo	2,81	2,20	2,31	2,54	2,34	2,29
y3(2)	Temp. topo	51,11	49,43	51,54	51,06	51,67	51,45
y3(3)	Vazão Fundo	2,75	2,65	2,63	2,01	2,36	2,87
y3(4)	Temp.	125,96	112,27	117,29	118,45	119,78	116,4
Perda IP	< 4	0,63	1,01	1,01	0,65	0,65	1,51

Notas: Perda IP – perda de IP no reator

Tabela E.2 - Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 1 Ótimo da Seção de Preparação de Carga

Variáveis		Simulação	Caso 1				
Notação	Descrição		a	b	c	d	e
x4(1)	Vazão carga	7,88	6,29	6,29	6,29	6,29	6,29
x4(2)	% IP Carga	30,63	30,76	30,76	30,76	30,76	30,76
x4(3)	% CPD Carga	3,34	1,36	1,36	1,36	1,36	1,36
x4(4)	% 2M2B Carga	0,59	0,79	0,79	0,79	0,79	0,79
x4(5)	Vazão Solvente	56,1	57,7	57,2	62,8	62,9	68,0
x4(6)	Refluxo	12,75	10,62	10,62	10,62	10,62	12,97
x4(7)	Vapor Referv.	3,98	4,59	3,49	2,40	2,40	3,49
y4(1)	Vazão topo I	3,53	2,74	3,10	2,43	2,42	2,27
y4(2)	Vazão topo II	3,49	3,41	3,72	3,87	3,87	3,84
y4(3)	% i-C5 topo	29,77	37,70	35,58	27,61	27,60	29,04
y4(4)	% n-C5 topo	38,81	39,37	40,18	37,87	37,84	38,09
y4(5)	Vazão fundo	53,53	57,75	56,09	67,04	67,12	69,31
y4(6)	Temp. Fundo	85,96	90,28	87,58	88,09	88,10	88,30
x5(1)	Vazão Carga I	53,53	57,75	56,09	67,04	67,12	69,31
x5(2)	Temp. Carga I	85,96	90,28	87,58	88,09	88,10	88,30
x5(3)	Vazão Carga II	1,18	1,00	1,00	1,00	1,00	1,00
x5(4)	Temp. Carga II	92,33	90,89	90,89	90,89	90,89	90,89
x5(5)	Vazão Solvente	14,02	14,30	14,30	13,00	13,00	15,60
x5(6)	Refluxo	15,01	14,74	14,74	13,57	13,57	14,74
x5(7)	Vapor Referv.	5,54	4,82	4,82	4,82	4,82	5,74
y5(1)	Vazão Topo	2,51	2,26	2,44	2,40	2,40	2,70
y5(2)	% Butino-2 topo	1,43	1,72	1,66	1,75	1,75	1,67
y5(3)	% CPD topo	4,12	1,43	1,29	0,78	0,77	1,94
y5(4)	Retirada Lateral	1,72	1,57	1,54	1,52	1,52	1,56
y5(5)	Temp. # 55	98,7	98,22	98,11	98,48	98,49	98,98
y5(6)	Vazão Fundo	68,25	69,34	69,03	73,29	73,37	80,39
y5(7)	Temp. Fundo	102,6	102,08	101,76	101,84	101,84	102,93
y5(8)	% H2O no solv.	14,71	14,37	15,52	20,77	20,81	22,95
x6(1)	Temp. Topo	63,34	63,34	63,34	63,34	63,34	63,34
x6(2)	Temp. Vaso Topo	28,48	28,48	28,48	28,48	28,48	28,48
x6(3)	% IP Topo	6,71	6,71	6,71	6,71	6,71	6,71
x6(4)	% Butino-2 Topo	17,65	17,65	17,65	17,65	17,65	17,65
x6(5)	% CPD Topo	47,83	47,83	47,83	47,83	47,83	47,83
x6(6)	Vazão Fundo	1,18	1,00	1	1	1	1
x6(7)	Temp. Fundo	92,33	90,89	90,89	90,89	90,89	90,89
y6(1)	Vazão de Carga	1,77	1,57	1,57	1,57	1,57	1,57
y6(2)	Refluxo	1,5	1,4	1,4	1,4	1,4	1,4
x7(1)	Vazão carga I	7,82	6,93	7,58	7,06	7,06	6,88
x7(2)	Vazão carga II	2,51	2,26	2,44	2,40	2,40	2,70
x7(3)	% Butino-2 carga	1,43	1,72	1,66	1,75	1,75	1,67
x7(4)	% CPD carga	4,12	1,43	1,29	0,78	0,77	1,94
x7(5)	Vazão água I	2,45	2,69	2,69	2,69	1,90	2,29
x7(6)	Vazão água II	2,5	2,5	2,48	2,48	1,3	2,48

Tabela E.2 (cont.) - Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 1 Ótimo da Seção de Preparação de Carga

Variáveis		Caso 1					
Notação	Descrição	a	b	c	d	e	f
y7(1)	Vazão topo	2,27	2,11	2,24	2,21	2,12	2,47
y7(2)	Temp. topo	29,23	29,42	29,53	29,51	28,66	29,39
y7(3)	Vazão fundo	6,15	6,12	6,12	6,10	3,92	5,96
x8(1)	Vazão carga	6,15	6,12	6,12	6,10	3,92	5,96
x8(2)	Temp. carga	78,26	78,26	78,26	78,26	78,26	78,26
x8(3)	Refluxo	0,79	1,05	0,6	0,6	0,6	1,05
x8(4)	Vapor Referv.	1,63	1,26	1,43	1,26	1,26	1,6
y8(1)	Vazão topo	1,67	1,26	1,11	1,09	1,12	1,77
y8(2)	Temp. topo	100,63	100,63	100,63	100,63	100,75	100,65
y8(3)	% H2O topo	20,3	20,24	20,57	20,41	21,53	20,40
y8(4)	Temp. # 10	105,45	105,48	105,43	105,43	105,52	105,6
y5(7)-y5(5)	ΔT (3 a 4)		3,86	3,65	3,36	3,35	3,95
x4(5)/x4(1)	solv/carga (7,8)		9,17	9,09	9,97	9,99	10,80
x5(5)/y5(1)	solv/topo (5 a 6)		6,33	5,86	5,41	5,41	5,77
x7(5)/x7(2)	água/carga (1,0)		1,19	1,10	1,12	0,79	0,85
x7(6)/x7(1)	água/carga (0,4)		0,36	0,33	0,35	0,18	0,36
Casos que satisfazem as restrições			17658				

Tabela E.3 - Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 2 Ótimo da Seção de Preparação de Carga

Variáveis		Simulação	Caso 2				
Notação	Descrição		a	b	c	d	e
x4(1)	Vazão carga	7,88	5,64	5,641	5,641	5,641	5,641
x4(2)	% IP Carga	30,63	33,16	33,16	33,16	33,16	33,16
x4(3)	% CPD Carga	3,34	0,90	0,90	0,90	0,90	0,90
x4(4)	% 2M2B Carga	0,59	0,85	0,85	0,85	0,85	0,85
x4(5)	Vazão Solvente	56,1	61,2	59,30	64,08	64,12	67,26
x4(6)	Refluxo	12,75	10,62	15,32	10,62	10,62	12,97
x4(7)	Vapor Referv.	3,98	4,59	2,40	2,40	2,40	2,40
y4(1)	Vazão topo I	3,53	1,96	1,62	2,20	2,20	2,04
y4(2)	Vazão topo II	3,49	3,66	4,80	3,80	3,80	4,01
y4(3)	% i-C5 topo	29,77	39,47	32,88	27,93	27,91	27,12
y4(4)	% n-C5 topo	38,81	37,10	37,25	37,24	37,24	36,68
y4(5)	Vazão fundo	53,53	61,78	61,69	67,84	67,87	69,65
y4(6)	Temp. Fundo	85,96	91,73	86,36	88,80	88,80	87,41
x5(1)	Vazão Carga I	53,53	61,78	61,69	67,84	67,87	69,65
x5(2)	Temp. Carga I	85,96	91,73	86,36	88,80	88,80	87,41
x5(3)	Vazão Carga II	1,18	1,00	1,00	1,00	1,00	1,00
x5(4)	Temp. Carga II	92,33	90,89	90,89	90,89	90,89	90,89
x5(5)	Vazão Solvente	14,02	13,00	14,30	13,00	13,00	15,60
x5(6)	Refluxo	15,01	14,74	13,57	13,57	13,57	14,74
x5(7)	Vapor Referv.	5,54	5,28	4,82	4,82	4,82	5,74
y5(1)	Vazão Topo	2,51	2,25	2,42	2,38	2,38	2,70
y5(2)	% Butino-2 topo	1,43	1,74	1,75	1,73	1,73	1,73
y5(3)	% CPD topo	4,12	0,95	1,92	0,73	0,73	1,09
y5(4)	Retirada Lateral	1,72	1,56	1,64	1,50	1,50	1,59
y5(5)	Temp. # 55	98,7	98,55	98,22	98,70	98,70	99,23
y5(6)	Vazão Fundo	68,25	71,19	70,66	74,58	74,60	80,36
y5(7)	Temp. Fundo	102,6	102,58	101,91	101,92	101,92	103,08
y5(8)	% H2O no solv.	14,71	14,63	14,59	22,65	22,64	22,08
x6(1)	Temp. Topo	63,34	63,34	63,34	63,34	63,34	63,34
x6(2)	Temp. Vaso Topo	28,48	28,48	28,48	28,48	28,48	28,48
x6(3)	% IP Topo	6,71	6,71	6,71	6,71	6,71	6,71
x6(4)	% Butino-2 Topo	17,65	17,65	17,65	17,65	17,65	17,65
x6(5)	% CPD Topo	47,83	47,83	47,83	47,83	47,83	47,83
x6(6)	Vazão Fundo	1,18	1,00	1	1	1	1
x6(7)	Temp. Fundo	92,33	90,89	90,89	90,89	90,89	90,89
y6(1)	Vazão de Carga	1,77	1,57	1,57	1,57	1,57	1,57
y6(2)	Refluxo	1,5	1,4	1,4	1,4	1,4	1,4
x7(1)	Vazão carga I	7,82	6,39	7,20	6,78	6,77	6,83
x7(2)	Vazão carga II	2,51	2,25	2,42	2,38	2,38	2,70
x7(3)	% Butino-2 carga	1,43	1,74	1,75	1,73	1,73	1,73
x7(4)	% CPD carga	4,12	0,95	1,92	0,73	0,73	1,09
x7(5)	Vazão água I	2,45	2,29	1,90	2,69	1,90	1,90
x7(6)	Vazão água II	2,5	2,5	2,48	2,48	1,3	2,48

Tabela E.3 (cont.)- Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 2 Ótimo da Seção de Preparação de Carga

Variáveis		Simulação	Caso 2				
Notação	Descrição		a	b	c	d	e
y7(1)	Vazão topo	2,27	2,04	2,18	2,19	2,10	2,48
y7(2)	Temp. topo	29,23	29,14	29,30	29,48	28,57	29,02
y7(3)	Vazão fundo	6,15	5,87	5,57	6,10	3,92	5,46
x8(1)	Vazão carga	6,15	5,87	5,57	6,10	3,92	5,46
x8(2)	Temp. carga	78,26	78,26	78,26	78,26	78,26	78,26
x8(3)	Refluxo	0,79	1,05	1,5	0,6	0,6	0,6
x8(4)	Vapor Referv.	1,63	1,43	1,6	1,26	1,26	1,26
y8(1)	Vazão topo	1,67	1,57	1,54	1,09	1,12	1,10
y8(2)	Temp. topo	100,63	100,63	101,44	100,63	100,75	100,63
y8(3)	% H2O topo	20,3	20,24	24,04	20,41	21,50	20,31
y8(4)	Temp. # 10	105,45	105,51	107,37	105,43	105,52	105,43
y5(7)-y5(5)	ΔT (3 a 4)		4,04	3,69	3,22	3,22	3,85
x4(5)/x4(1)	solv/carga (7,8)		10,84	10,51	11,36	11,37	11,92
x5(5)/y5(1)	solv/topo (5 a 6)		5,79	5,90	5,46	5,46	5,78
x7(5)/x7(2)	água/carga (1,0)		1,02	0,78	1,13	0,80	0,70
x7(6)/x7(1)	água/carga (0,4)		0,39	0,34	0,37	0,19	0,36
Casos que satisfazem as restrições			12447				

Tabela E.4 - Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 3 Ótimo da Seção de Preparação de Carga

Variáveis		Simulação	Caso 3					
Notação	Descrição		a	b	c	d	e	f
x4(1)	Vazão carga	7,88	7,49	7,49	7,49	7,49	7,49	7,49
x4(2)	% IP Carga	30,63	30,83	30,83	30,83	30,83	30,83	30,83
x4(3)	% CPD Carga	3,34	2,00	2,00	2,00	2,00	2,00	2,00
x4(4)	% 2M2B Carga	0,59	0,96	0,96	0,96	0,96	0,96	0,96
x4(5)	Vazão Solvente	56,1	57,6	57,50	57,45	57,00	57,32	57,30
x4(6)	Refluxo	12,75	10,62	10,62	10,62	15,32	10,62	10,62
x4(7)	Vapor Referv.	3,98	2,40	2,40	2,40	2,40	2,40	2,40
y4(1)	Vazão topo I	3,53	3,76	3,76	3,76	3,66	3,76	3,76
y4(2)	Vazão topo II	3,49	4,14	4,14	4,13	3,09	4,12	4,12
y4(3)	% i-C5 topo	29,77	29,71	29,69	29,67	34,84	29,64	29,63
y4(4)	% n-C5 topo	38,81	41,142	41,14	41,13	41,44	41,12	41,12
y4(5)	Vazão fundo	53,53	53,78	53,76	53,75	55,48	53,72	53,71
y4(6)	Temp. Fundo	85,96	84,42	84,42	84,42	86,62	84,43	84,43
x5(1)	Vazão Carga I	53,53	53,78	53,76	53,75	55,48	53,72	53,71
x5(2)	Temp. Carga I	85,96	84,42	84,42	84,42	86,62	84,43	84,43
x5(3)	Vazão Carga II	1,18	1,00	1,00	1,00	1,00	1,00	1,00
x5(4)	Temp. Carga II	92,33	90,89	90,89	90,89	90,89	90,89	90,89
x5(5)	Vazão Solvente	14,02	13,00	13,00	13,00	14,30	13,00	13,00
x5(6)	Refluxo	15,01	14,74	14,74	14,74	14,74	14,74	14,74
x5(7)	Vapor Referv.	5,54	5,74	5,74	5,74	4,82	5,74	5,74
y5(1)	Vazão Topo	2,51	2,21	2,21	2,21	2,53	2,21	2,21
y5(2)	% Butino-2 topo	1,43	1,67	1,67	1,67	1,64	1,67	1,67
y5(3)	% CPD topo	4,12	1,39	1,39	1,39	1,95	1,38	1,38
y5(4)	Retirada Lateral	1,72	1,67	1,67	1,67	1,53	1,67	1,67
y5(5)	Temp. # 55	98,7	99,36	99,36	99,36	98,02	99,37	99,37
y5(6)	Vazão Fundo	68,25	67,81	67,81	67,81	68,80	67,81	67,81
y5(7)	Temp. Fundo	102,6	103,06	103,07	103,07	101,69	103,07	103,07
y5(8)	% H2O no solv.	14,71	14,49	14,49	14,49	16,01	14,48	14,48
x6(1)	Temp. Topo	63,34	63,34	63,34	63,34	63,34	63,34	63,34
x6(2)	Temp. Vaso Topo	28,48	28,48	28,48	28,48	28,48	28,48	28,48
x6(3)	% IP Topo	6,71	6,71	6,71	6,71	6,71	6,71	6,71
x6(4)	% Butino-2 Topo	17,65	17,65	17,65	17,65	17,65	17,65	17,65
x6(5)	% CPD Topo	47,83	47,83	47,83	47,83	47,83	47,83	47,83
x6(6)	Vazão Fundo	1,18	1,00	1,00	1,00	1,00	1,00	1,00
x6(7)	Temp. Fundo	92,33	90,89	90,89	90,89	90,89	90,89	90,89
y6(1)	Vazão de Carga	1,77	1,57	1,57	1,57	1,57	1,57	1,57
y6(2)	Refluxo	1,50	1,40	1,40	1,40	1,40	1,40	1,40
x7(1)	Vazão carga I	7,82	8,67	8,67	8,66	7,52	8,65	8,65
x7(2)	Vazão carga II	2,51	2,21	2,21	2,21	2,53	2,21	2,21
x7(3)	% Butino-2 carga	1,43	1,67	1,67	1,67	1,64	1,67	1,67
x7(4)	% CPD carga	4,12	1,39	1,39	1,39	1,95	1,38	1,38
x7(5)	Vazão água I	2,45	2,69	2,29	2,69	2,69	1,90	2,29
x7(6)	Vazão água II	2,50	1,90	2,48	2,48	1,89	1,30	2,48

Tabela E.4 (cont.) - Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 3 Ótimo da Seção de Preparação de Carga

Variáveis		Simulação	Caso 3					
Notação	Descrição		a	b	c	d	e	f
y7(1)	Vazão topo	2,27	2,17	2.07	2,15	2,30	2,06	2,07
y7(2)	Temp. topo	29,23	29,05	30.59	28,33	29,06	29,04	30,61
y7(3)	Vazão fundo	6,15	5,66	5.99	6,24	5,42	3,97	5,99
x8(1)	Vazão carga	6,15	5,66	5.99	6,24	5,42	3,97	5,99
x8(2)	Temp. carga	78,26	78,26	78.26	78,26	78,26	78,26	78,26
x8(3)	Refluxo	0,79	1,05	1.05	1,05	0,60	0,60	0,60
x8(4)	Vapor Referv.	1,63	1,26	1.26	1,26	1,26	1,26	1,26
y8(1)	Vazão topo	1,67	1,35	1.29	1,24	1,10	1,12	1,09
y8(2)	Temp. topo	100,63	100,63	100.63	100,63	100,63	100,71	100,63
y8(3)	% H2O topo	20,3	20,24	20.24	20,24	20,31	21,21	20,39
y8(4)	Temp. # 10	105,45	105,54	105.5	105,47	105,43	105,5	105,43
y5(7)-y5(5)	ΔT (3 a 4)		3,70	3.71	3,71	3,67	3,70	3,70
x4(5)/x4(1)	solv/carga (7,8)		7,68	7.67	7,67	7,61	7,65	7,65
x5(5)/y5(1)	solv/topo (5 a 6)		5,88	5.88	5,88	5,66	5,88	5,88
x7(5)/x7(2)	água/carga (1,0)		1,22	1.04	1,22	1,06	0,86	1,04
x7(6)/x7(1)	agua/carga (0,4)		0,22	0.29	0,29	0,25	0,15	0,29
Casos que satisfazem as restrições			4552					

Tabela E.5 - Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 4 Ótimo da Seção de Preparação de Carga

Variáveis		Simulação	Caso 4				
Notação	Descrição		a	b	c	d	e
x4(1)	Vazão carga	7,88	7,17	7,17	7,17	7,17	7,17
x4(2)	% IP Carga	30,63	29,71	29,71	29,71	29,71	29,71
x4(3)	% CPD Carga	3,34	2,07	2,07	2,07	2,07	2,07
x4(4)	% 2M2B Carga	0,59	0,88	0,88	0,88	0,88	0,88
x4(5)	Vazão Solvente	56,10	64,25	56,89	57,35	56,89	57,97
x4(6)	Refluxo	12,75	15,32	15,32	10,62	15,32	12,97
x4(7)	Vapor Referv.	3,98	3,49	3,49	2,40	3,49	2,40
y4(1)	Vazão topo I	3,53	2,69	3,06	3,72	3,06	3,66
y4(2)	Vazão topo II	3,49	4,28	3,44	4,02	3,44	3,34
y4(3)	% i-C5 topo	29,77	27,72	28,00	29,62	28,00	31,84
y4(4)	% n-C5 topo	38,81	37,48	37,53	41,03	37,53	41,22
y4(5)	Vazão fundo	53,53	65,42	55,16	53,92	55,16	55,70
y4(6)	Temp. Fundo	85,96	87,75	86,73	84,52	86,73	85,42
x5(1)	Vazão Carga I	53,53	65,42	55,16	53,92	55,16	55,70
x5(2)	Temp. Carga I	85,96	87,75	86,73	84,52	86,73	85,42
x5(3)	Vazão Carga II	1,18	1,00	1,00	1,00	1,00	1,00
x5(4)	Temp. Carga II	92,33	90,89	90,89	90,89	90,89	90,89
x5(5)	Vazão Solvente	14,02	15,60	14,30	13,00	14,30	13,00
x5(6)	Refluxo	15,01	14,74	14,74	14,74	14,74	14,74
x5(7)	Vapor Referv.	5,54	4,82	4,82	5,74	4,82	5,28
y5(1)	Vazão Topo	2,51	2,65	2,51	2,21	2,51	2,28
y5(2)	% Butino-2 topo	1,43	1,71	1,62	1,67	1,62	1,50
y5(3)	% CPD topo	4,12	1,36	1,93	1,32	1,93	1,81
y5(4)	Retirada Lateral	1,72	1,67	1,53	1,66	1,53	1,57
y5(5)	Temp. # 55	98,70	98,22	98,04	99,38	98,04	98,41
y5(6)	Vazão Fundo	68,25	77,19	68,69	67,84	68,69	68,09
y5(7)	Temp. Fundo	102,60	102,17	101,70	103,07	101,70	101,86
y5(8)	% H2O no solv.	14,71	14,46	16,18	14,55	16,18	18,01
x6(1)	Temp. Topo	63,34	63,34	63,34	63,34	63,34	63,34
x6(2)	Temp. Vaso Topo	28,48	28,48	28,48	28,48	28,48	28,48
x6(3)	% IP Topo	6,71	6,71	6,71	6,71	6,71	6,71
x6(4)	% Butino-2 Topo	17,65	17,65	17,65	17,65	17,65	17,65
x6(5)	% CPD Topo	47,83	47,83	47,83	47,83	47,83	47,83
x6(6)	Vazão Fundo	1,18	1,00	1,00	1,00	1,00	1,00
x6(7)	Temp. Fundo	92,33	90,89	90,89	90,89	90,89	90,89
y6(1)	Vazão de Carga	1,77	1,57	1,57	1,57	1,57	1,57
y6(2)	Refluxo	1,50	1,40	1,40	1,40	1,40	1,40
x7(1)	Vazão carga I	7,82	7,74	7,27	8,51	7,27	7,77
x7(2)	Vazão carga II	2,51	2,65	2,51	2,21	2,51	2,28
x7(3)	% Butino-2 carga	1,43	1,71	1,62	1,67	1,62	1,50
x7(4)	% CPD carga	4,12	1,36	1,93	1,32	1,93	1,81
x7(5)	Vazão água I	2,45	2,69	2,29	2,69	2,69	2,29
x7(6)	Vazão água II	2,50	2,48	2,48	1,30	2,48	2,48

Tabela E.5 (cont.)- Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 4 Ótimo da Seção de Preparação de Carga

Variáveis		Simulação	Caso 4				
Notação	Descrição		a	b	c	d	e
y7(1)	Vazão topo	2,27	2,38	2,28	2,17	2,28	2,07
y7(2)	Temp. topo	29,23	29,65	29,31	28,52	29,56	29,38
y7(3)	Vazão fundo	6,15	6,13	5,93	4,55	6,16	5,93
x8(1)	Vazão carga	6,15	6,13	5,93	4,55	6,16	5,93
x8(2)	Temp. carga	78,26	78,26	78,26	78,26	78,26	78,26
x8(3)	Refluxo	0,79	1,05	0,60	0,60	0,60	1,50
x8(4)	Vapor Referv.	1,63	1,26	1,26	1,26	1,26	1,26
y8(1)	Vazão topo	1,67	1,26	1,09	1,11	1,09	1,48
y8(2)	Temp. topo	100,63	100,63	100,63	100,63	100,63	100,66
y8(3)	% H2O topo	20,30	20,24	20,37	20,29	20,42	20,26
y8(4)	Temp. # 10	105,45	105,48	105,43	105,43	105,43	108,67
y5(7)-y5(5)	ΔT (3 a 4)		3,95	3,66	3,69	3,66	3,45
x4(5)/x4(1)	solv/carga (7,8)		8,97	7,94	8,00	7,94	8,09
x5(5)/y5(1)	solv/topo (5 a 6)		5,88	5,69	5,88	5,69	5,70
x7(5)/x7(2)	água/carga (1,0)		1,01	0,91	1,22	1,07	1,01
x7(6)/x7(1)	agua/carga (0,4)		0,32	0,34	0,15	0,34	0,32
Casos que satisfazem as restrições			7380				

Tabela E.6 - Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 5 Ótimo da Seção de Preparação de Carga

Variável		Simulação	Caso 5						
Notação	Descrição		a	b	c	d	e	f	g
x4(1)	Vazão carga	7,88	7,46	7,46	7,46	7,46	7,46	7,46	7,46
x4(2)	% IP Carga	30,63	32,66	32,66	32,66	32,66	32,66	32,66	32,66
x4(3)	% CPD Carga	3,34	1,93	1,93	1,93	1,93	1,93	1,93	1,93
x4(4)	% 2M2B Carga	0,59	1,10	1,10	1,10	1,10	1,10	1,10	1,10
x4(5)	Vazão Solvente	56,10	57,61	57,68	56,93	57,40	57,42	57,40	57,80
x4(6)	Refluxo	12,75	10,62	10,62	12,97	10,62	10,62	10,62	15,32
x4(7)	Vapor Referv.	3,98	4,59	4,59	4,59	2,40	2,40	2,40	2,40
y4(1)	Vazão topo I	3,53	3,58	3,58	3,32	3,68	3,68	3,68	3,53
y4(2)	Vazão topo II	3,49	3,62	3,62	3,33	3,89	3,89	3,89	3,06
y4(3)	% i-C5 topo	29,77	30,62	30,54	34,10	28,99	29,00	28,99	34,22
y4(4)	% n-C5 topo	38,81	40,79	40,79	40,19	40,97	40,97	40,97	41,42
y4(5)	Vazão fundo	53,53	53,89	53,90	55,27	54,47	54,48	54,47	56,75
y4(6)	Temp. Fundo	85,96	85,09	85,09	86,81	84,52	84,52	84,52	86,14
x5(1)	Vazão Carga I	53,53	53,89	53,90	55,27	54,47	54,48	54,47	56,75
x5(2)	Temp. Carga I	85,96	85,09	85,09	86,81	84,52	84,52	84,52	86,14
x5(3)	Vazão Carga II	1,18	1,00	1,00	1,00	1,00	1,00	1,00	1,00
x5(4)	Temp. Carga II	92,33	90,89	90,89	90,89	90,89	90,89	90,89	90,89
x5(5)	Vazão Solvente	14,02	13,00	13,00	14,30	13,00	13,00	13,00	14,30
x5(6)	Refluxo	15,01	14,74	14,74	14,74	14,74	14,74	14,74	14,74
x5(7)	Vapor Referv.	5,54	5,74	5,74	4,82	5,74	5,74	5,74	4,82
y5(1)	Vazão Topo	2,51	2,22	2,22	2,51	2,21	2,21	2,21	2,56
y5(2)	% Butino-2 topo	1,43	1,65	1,65	1,63	1,66	1,66	1,66	1,72
y5(3)	% CPD topo	4,12	1,05	1,05	1,84	1,34	1,34	1,34	1,99
y5(4)	Retirada Lateral	1,72	1,66	1,66	1,53	1,66	1,66	1,66	1,55
y5(5)	Temp. # 55	98,70	99,51	99,51	98,04	99,33	99,33	99,33	97,98
y5(6)	Vazão Fundo	68,25	67,91	67,91	68,73	67,90	67,90	67,90	69,14
y5(7)	Temp. Fundo	102,60	103,19	103,18	101,71	102,98	102,98	102,98	101,68
y5(8)	% H2O no solv.	14,71	14,64	14,64	16,08	14,74	14,75	14,74	15,15
x6(1)	Temp. Topo	63,34	63,34	63,34	63,34	63,34	63,34	63,34	63,34
x6(2)	Temp. Vaso Topo	28,48	28,48	28,48	28,48	28,48	28,48	28,48	28,48
x6(3)	% IP Topo	6,71	6,71	6,71	6,71	6,71	6,71	6,71	6,71
x6(4)	% Butino-2 Topo	17,65	17,65	17,65	17,65	17,65	17,65	17,65	17,65
x6(5)	% CPD Topo	47,83	47,83	47,83	47,83	47,83	47,83	47,83	47,83
x6(6)	Vazão Fundo	1,18	1,00	1,00	1,00	1,00	1,00	1,00	1,00
x6(7)	Temp. Fundo	92,33	90,89	90,89	90,89	90,89	90,89	90,89	90,89
y6(1)	Vazão de Carga	1,77	1,57	1,57	1,57	1,57	1,57	1,57	1,57
y6(2)	Refluxo	1,50	1,4	1,4	1,4	1,4	1,4	1,4	1,4
x7(1)	Vazão carga I	7,82	7,97	7,97	7,43	8,34	8,34	8,34	7,36
x7(2)	Vazão carga II	2,51	2,22	2,22	2,51	2,21	2,21	2,21	2,56
x7(3)	% Butino-2 carga	1,43	1,65	1,65	1,63	1,66	1,66	1,66	1,72
x7(4)	% CPD carga	4,12	1,05	1,05	1,84	1,34	1,34	1,34	1,99
x7(5)	Vazão água I	2,45	2,29	2,69	2,29	2,29	1,90	2,69	2,69
x7(6)	Vazão água II	2,50	2,5	1,89	2,48	2,48	1,3	1,89	1,89

Tabela E.6 (cont.)- Resultados da Otimização da Seção de Destilação Extrativa e Regeneração de Solvente – Caso 6 Ótimo da Seção de Preparação de Carga

Variável		Simulação	Caso 5						
Notação	Descrição		a	b	c	d	e	f	g
y7(1)	Vazão topo	2,27	2,05	2,16	2,27	2,05	2,05	2,16	2,33
y7(2)	Temp. topo	29,23	30,37	28,62	29,34	30,64	29,02	28,57	29,07
y7(3)	Vazão fundo	6,15	5,92	5,55	5,92	5,96	3,97	5,63	5,39
x8(1)	Vazão carga	6,15	5,92	5,55	5,92	5,96	3,97	5,63	5,39
x8(2)	Temp. carga	78,26	78,26	78,26	78,26	78,26	78,26	78,26	78,26
x8(3)	Refluxo	0,79	1,05	1,05	0,6	0,6	0,6	0,6	1,05
x8(4)	Vapor Referv.	1,63	1,26	1,26	1,26	1,26	1,26	1,26	1,43
y8(1)	Vazão topo	1,67	1,30	1,37	1,09	1,09	1,12	1,10	1,56
y8(2)	Temp. topo	100,63	100,63	100,63	100,63	100,63	100,71	100,63	100,69
y8(3)	% H2O topo	20,30	20,24	20,24	20,37	20,38	21,21	20,33	20,38
y8(4)	Temp. # 10	105,45	105,51	105,57	105,43	105,43	105,5	105,43	105,74
y5(7)-y5(5)	DT (3 a 4)		3,68	3,68	3,67	3,65	3,65	3,65	3,70
x4(5)/x4(1)	solv/carga (7,8)		7,72	7,73	7,63	7,69	7,69	7,69	7,74
x5(5)/y5(1)	solv/topo (5 a 6)		5,85	5,85	5,70	5,88	5,88	5,88	5,58
x7(5)/x7(2)	água/carga (1,0)		1,03	1,21	0,91	1,04	0,86	1,21	1,05
x7(6)/x7(1)	agua/carga (0,4)		0,31	0,24	0,33	0,30	0,16	0,23	0,26
Casos que satisfazem as restrições			4837						

Tabela E.7 - Resultados da Otimização da Seção de Fracionamento

Notação	Variável	Simulação	Caso 3b			Caso 4b			Caso 4d			Caso 5a		
			3b-1	3b-2	3b-3	4b-1	4b-2	4b-3	4d-1	4d-2	4d-3	5a-1	5a-2	5a-3
x9(1)	Vazão carga	2,27	2,07	2,07	2,07	2,28	2,28	2,28	2,28	2,28	2,28	2,05	2,05	2,05
x9(2)	Temp carga	29,23	30,60	30,60	30,60	29,31	29,31	29,31	29,56	29,56	29,56	30,38	30,38	30,38
x9(3)	Refluxo	13,05	9,36	12,75	9,36	9,36	12,75	9,36	9,36	12,75	9,36	12,75	9,36	9,36
x9(4)	Vazão solv. referv.	53,92	38,26	51,12	38,26	38,26	51,12	38,26	38,26	51,12	38,26	51,12	38,26	38,26
x9(5)	Temp entra solv referv	102,28	102,05	103,74	103,18	102,05	102,61	103,18	102,61	103,18	102,05	103,74	102,61	102,61
x9(6)	Temp sai solv referv	70,68	70,68	70,68	70,68	70,68	70,68	70,68	70,68	70,68	70,68	70,68	70,68	70,68
y9(1)	Temp vaso	39,98	40,14	41,99	40,26	39,40	41,57	40,53	40,65	41,69	40,92	39,66	41,91	39,99
y9(2)	Temp fundo	62,09	63,62	61,72	62,17	63,89	61,37	62,66	63,34	61,51	62,85	63,58	61,71	62,63
y9(3)	Vazão fundo	1,90	2,64	1,78	2,57	2,69	1,88	2,64	2,68	1,87	2,65	2,64	1,76	2,61
x10(1)	Temp carga	62,09	63,62	61,72	62,17	63,89	61,37	62,66	63,34	61,51	62,85	63,58	61,71	62,63
x10(2)	Vazão carga	1,90	2,64	1,78	2,57	2,69	1,88	2,64	2,68	1,87	2,65	2,64	1,76	2,61
x10(3)	Refluxo	13,22	14,10	14,10	11,39	14,10	14,10	11,39	14,10	14,10	11,39	14,10	14,10	11,39
x10(4)	Vazão cond. Referv.	13,05	13,05	13,05	13,05	13,05	13,05	13,05	13,05	13,05	13,05	13,05	13,05	13,05
x10(5)	Vazão solv. Referv.	18,16	18,16	18,16	18,16	18,16	18,16	18,16	18,16	18,16	18,16	18,16	18,16	18,16
x10(6)	Temp sai solv referv	64,19	64,19	64,19	64,19	64,19	64,19	64,19	64,19	64,19	64,19	64,19	64,19	64,19
y10(1)	Vazão topo	1,87	2,48	1,91	2,34	2,49	1,97	2,37	2,49	1,95	2,35	2,47	1,91	2,35
y10(2)	CPD topo	0,37	0,15	0,17	0,16	0,11	0,14	0,14	0,19	0,17	0,14	0,16	0,17	0,14
y10(3)	Temp fundo	52,61	53,27	53,43	51,60	53,35	53,06	52,40	53,26	53,21	52,56	53,26	53,43	52,26

ANEXO F – ANÁLISE DE COMPONENTES PRINCIPAIS (PCA)

Os resultados mostrados são referentes à análise de PCA para redução do número de variáveis para predição de azeótropos, empregando modelos de redes neurais.

Resultado do Programa SAS para Analise de Componentes Principais para Selecao de Variaveis para Predicao de Azeotropos

The PRINCOM Procedure

Observations 404
Variables 19

Simple Statistics

	MM1	Teb1	Tc1	Pc1	Zc1	w1	alfa1	beta1	pl1	MM2	Teb2	Tc2	Pc2	Zc2	w2
Mean	70.47419	343.7074257	523.5564356	51.59777	0.256168	0.362077	0.14297	0.271386	0.374381	79.91033	374.1634	562.9889	51.09356	0.255646	0.402683
STD	24.10061	38.2810923	56.2010591	34.25291	0.022652	0.165827	0.205248	0.227584	0.254234	25.31754	26.05889	43.90893	40.06781	0.018774	0.156087

Correlation Matrix

	MM1	Teb1	Tc1	Pc1	Zc1	w1	alfa1	beta1	pl1	MM2	Teb2	Tc2	Pc2	Zc2	w2
MM1	1														
Teb1	0.3275	1													
Tc1	0.2475	0.9095	1												
Pc1	-0.5662	0.11	0.3818	1											
Zc1	0.5977	-0.2475	-0.2184	-0.295	1										
w1	-0.2641	0.4697	0.1925	0.1093	-0.3943	1									
alfa1	-0.553	0.2713	0.2983	0.7635	-0.3419	0.6699	1								
beta1	-0.3888	0.3391	0.1722	0.1994	-0.5417	0.7781	0.5333	1							
pl1	-0.1303	0.3522	0.4051	0.1847	-0.5646	0.1979	0.1416	0.6267	1						
MM2	0.0172	0.3258	0.2844	0.1024	-0.126	0.2307	0.1879	0.1593	0.1125	1					
Teb2	0.111	0.5604	0.4933	0.109	-0.1594	0.3373	0.2338	0.224	0.1921	0.5438	1				
Tc2	0.096	0.4494	0.3789	0.0636	-0.1296	0.3035	0.1929	0.2079	0.1524	0.7898	0.7898	1			
Pc2	0.0116	-0.0577	-0.0663	-0.0594	0.0187	-0.0235	-0.0522	-0.0061	-0.0137	-0.6885	-0.1116	0.355	1		
Zc2	-0.001	0.1314	0.12	0.0606	-0.0765	0.092	0.0908	0.0717	0.0626	0.6319	0.1678	0.0551	-0.3313	1	
w2	0.0379	-0.0088	0.0217	-0.0112	0.0278	-0.1044	-0.0825	-0.0613	-0.0017	-0.359	0.011	-0.3913	-0.0058	-0.1556	1
alfa2	0.0518	-0.0732	-0.0591	-0.0633	0.0677	-0.1006	-0.1069	-0.0681	-0.0317	-0.7443	-0.1141	-0.0188	0.7243	-0.2919	0.636
beta2	0.0358	-0.0267	0.0062	-0.0102	0.0563	-0.1144	-0.0892	-0.0992	-0.0551	-0.5313	-0.1332	-0.3375	0.1526	-0.3243	0.7894
pl2	0.0636	0.1088	0.105	0.0038	0.0316	0.014	-0.0003	-0.0117	-0.032	-0.1669	0.1621	0.2913	0.1096	-0.47	-0.0761
Yexp	-0.1	0.1716	0.1995	0.1822	-0.0958	0.089	0.1689	0.0514	-0.0025	-0.2729	-0.225	-0.1136	0.2219	-0.2165	0.0205

The PRINCOM Procedure

Eigenvalues of the Correlation Matrix			
Eigenvalue	Difference	Proportion	Cumulative
1	4.764227	1.55652858	0.2507
2	3.207699	0.66143098	0.1688
3	2.546268	0.79575198	0.134
4	1.750516	0.3517278	0.0921
5	1.398788	0.04830365	0.0736
6	1.350484	0.16113514	0.0711
7	1.189349	0.1805411	0.0626
8	1.028808	0.43286068	0.0541
9	0.595947	0.11976138	0.0314
10	0.476186	0.11278489	0.0251
11	0.363401	0.25751079	0.0191
12	0.10589	0.0300691	0.0056
13	0.075821	0.01898841	0.004
14	0.056833	0.00948514	0.003
15	0.047348	0.0247795	0.0025
16	0.022568	0.00749559	0.0012
17	0.015073	0.01208768	0.0008
18	0.002985	0.00117763	0.0002
19	0.001807	0.0001	1

Eigenvectors

	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8	Prin8	Prin9	Prin10	Prin11	Prin12	Prin13	Prin14	Prin15
MM1	-0.12022	-0.186552	0.490438	0.109868	-0.09134	0.156954	-0.12857	0.178211	0.277391	-0.04036	0.008389	0.492388	0.152775	0.517489	0.012433	
Teb1	0.297639	0.106191	0.386747	0.178167	0.014275	0.192388	-0.14345	0.143244	0.005691	0.129326	-0.28994	-0.23359	-0.0134	-0.08858	-0.01956	
Tc1	0.271321	0.117867	0.324526	0.234193	0.194291	0.275135	-0.13316	-0.16298	0.120424	0.068502	-0.216	-0.24924	-0.03292	-0.17192	-0.00666	
Pc1	0.199705	0.189719	-0.230781	0.068149	0.487919	0.128639	0.17045	-0.31716	0.260707	-0.08182	0.050183	-0.02953	0.087875	0.38987	-0.03592	
Zc1	-0.24933	-0.210609	0.207628	-0.02637	0.233685	0.086302	0.172059	0.297391	0.471663	-0.22761	0.421353	-0.28581	-0.12436	-0.3309	-0.00045	
w1	0.303861	0.169916	-0.059952	-0.08104	-0.15009	-0.11197	0.061176	0.602632	0.017631	0.060101	-0.14694	0.178737	-0.02406	-0.17083	0.013944	
alfa1	0.294912	0.215168	-0.224837	-0.02906	0.31094	0.021853	0.22185	0.210914	0.253069	-0.04423	-0.05001	0.436459	-0.00251	-0.15186	0.06667	
beta1	0.300728	0.218638	-0.145578	-0.0391	-0.33609	-0.14132	-0.1281	0.230956	0.231001	-0.07527	0.281046	-0.46512	0.091403	0.439785	-0.07427	
pi1	0.227823	0.166407	-0.003405	0.123119	-0.38277	-0.00888	-0.34354	-0.40898	0.286326	-0.17407	0.326581	0.307307	-0.0914	-0.34208	0.061738	
Teb2	0.311715	-0.35981	0.009394	0.157217	0.052723	-0.08732	0.070509	0.006133	-0.16418	-0.00503	0.166923	-0.06209	0.073127	0.056551	0.808152	
Tc2	0.288806	-0.014094	0.33728	0.009834	0.036791	-0.22766	0.303222	-0.07689	-0.31066	-0.33193	0.167993	0.0474	-0.13899	0.046184	-0.18244	
Pc2	0.246528	0.023204	0.341516	-0.40035	-0.01236	-0.06463	0.258985	-0.14931	-0.14185	-0.03075	0.144914	0.042288	-0.07477	0.028156	-0.18532	
w2	-0.12838	0.314387	-0.146473	-0.42778	-0.18028	0.223425	0.226133	-0.12185	0.103018	0.192439	-0.03073	-0.02158	-0.22562	0.049757	0.374478	
Zc2	0.173786	-0.268925	-0.081406	0.245916	-0.18847	0.168423	0.370448	-0.05925	0.052571	0.660918	0.296137	0.052837	0.080482	-0.05853	-0.03179	
alfa2	-0.16319	0.272285	0.062663	0.52242	-0.04918	-0.1544	0.215413	0.045547	-0.16442	-0.22678	0.052679	0.023156	0.253643	-0.07951	-0.03179	
beta2	-0.21213	0.392531	0.137659	0.031602	-0.19418	0.089429	0.349016	-0.05174	0.023766	0.01537	0.012746	-0.0843	0.47408	-0.07997	0.195583	
pi2	-0.19935	0.35215	-0.10373	0.343637	0.114985	-0.28448	-0.00244	0.059732	-0.02848	0.298333	0.152187	0.040082	-0.52279	0.177609	0.11282	
Yexp	-0.01233	0.138868	0.228598	-0.22314	0.354435	-0.47853	-0.32827	-0.02439	0.018648	0.388523	0.217677	0.021733	0.411827	-0.11853	0.015409	
	-0.0061	0.216582	-0.056548	-0.03984	0.175233	-0.28892	0.219901	-0.48247	-0.01738	0.484048	0.0269	0.084819	0.0269	0.018219	-0.01522	

LISTA DE REFERÊNCIAS

- **Redes Neurais:**

- Assaf, E.M., Giordano, R.C. & Nascimento, C.A.O., "Thermal Runway of Ethylene Oxidation Reactors: Prevision Through Neuronal Networks", *Chem. Engng. Sci.*, 1996, 51, pp. 3107-3112.
- Becraft, W.R. & Lee, P.L., "An Integrated Neural Network/Expert System Approach for Fault Diagnosis", *Comp. & Chem. Engng*, 1993, 17, pp. 1001-1014.
- Bhat, N.V. & McAvoy, T., "Use of Neural Nets for Dynamic Modeling and Control of Chemical Process Systems", *Computers & Chemical Engineering*, 1990, v. 14(4/5), pp. 573-583.
- Bhat, N.V., Minderman Jr., P.A., McAvoy, T. & Wang, N.S., "Modeling Chemical Process Systems via Neural Computation", *IEEE Control Systems Magazine*, 1990, v.4, pp. 24-30.
- Bhat, N. & McAvoy, T.J., "Use of Neural Nets for Dynamic Modelling and Control of Chemical Process Systems". *Proc. Amer. Cont. Conf.*, 1989, Pittsburg, v.2, pp. 1342-1347.
- Bhat, N.; Mindermam, P.; McAvoy, T.J., Use of Neural Nets for Modelling of Chemical Process Systems. IFAC Symp. Dycord +89, Maastricht, The Netherlands, 1989, Aug., pp. 147-153.
- Bhat, N.; Minderman, P.; McAvoy, T.; Wang, N., Modelling Chemcial Process Systems via Neural Computational, Proc. 3rd. Int. Symp. Control for Profit, Newcastle-upon-Tyne, 1989.
- Blum,J.; Villard, P.; Leuba, A.; Himmelblau, D.M., Practical Issues in Applying Artificial Neural Networks for Identification in Model Based Predicitive Control, *AIChE Annual Meeting*, 1992, Miami.
- Bülau, H. C., Ulrich, J., Guardani, R., Nascimento, C. A. O., "Application of Neural Networks to Data from a Melt Crystalization Process for the Detection of Outliers". Em *Proceedings of AIDA, International Seminar on Advances in Data Analysis*, 1999 – Washington, DC

- Bulsari, A.B., Applications of Artificial Neural Networks in Process Engineering, *Journal of Systems Engineering*, 1994, 4, pp. 131-170.
- Chan, W.M. & Nascimento, C.A.O., "Use of Neural Networks for Modeling of Olefin Polymerization in High Pressure Tubular Reactors", *Journal of Applied Polymer Science*, 1994, 53, 1277-1289.
- Chen, J.; Bruns, D.D., Recursive Nonlinear System Identification Using an adaptive Wave ARX Network, *Ind. Engng. Chem. Res.*, 1996, 35, pp. 2782-2789.
- Crowe, E.R.; Vassiliadis, C.A., Artificial Intelligence: Starting to Realize its Practical Promise, *Chem Engng. Progr.*, 1995, Jan., pp. 22-31.
- Cubillos, F. A.; Alvarez, P.I.; Pinto, J.C.; Lima, E.L., Hybrid-Neural Modeling for Particulate Solid Drying Process, *Powder Technology*, 1996, 87, pp. 153-160.
- Cybenko, G., Approximations by Superpositions of a Sigmoidal Function, *Math. Control Signal Syst.*, 1989, 2, pp. 303-314.
- Dayhoff, J.E., *Neural Network Architectures – An Introduction*, Van Nostrand Reinhold, New York, 1990.
- Di Massimo, C.; Willis, M.J.; Montague, G. A.; Kambhampati, C.; Hofland, A.G.; Tham, M.T.; Morris, A.J., On the Applicability of Neural Networks in Chemical Process Control, *AIChE Annual Meeting*, 1990, Chicago.
- Doherty, S.K.; Gomm, J.B.; Williams, D., Experimental Design Considerations for non-linear System Identification Using Neural Networks, *Computers & Chem. Engng.*, 1997, 21, pp. 327-346.
- Dong, D.; McAvoy, T.J.; Zafiriou, E., Batch-to-Batch Optimization Using Neural Network Models, *Ind. Engng. Chem. Res.*, 1996, 35, pp. 2269-2276.
- Epping, W.J.M.; Nitters, G., A Neural Network for Analysis and Improvement of Gas Well Production, *Proc. of Summer Computer Simulation Conf.*, 1990, pp. 329-334, Calgary, Canada.
- Feldman, J.A.; Ballard, D.H., Connectionist Models and their Properties, *Cognitive Sci.*, 1982, 6, pp. 205-254.
- Gokhale, V.; Hurowitz, S.; Riggs, J.B., Comparison of Advanced Distillation Control Techniques for a Propylene/Propane Splitter, *Ind. Engng. Chem. Res.*, 1995, 34, pp. 4413-4419.

- Haykin, S., *Neural Networks – A Comprehensive Foundation*, Macmillan College Publishing Co., New York, 1994.
- Hebb, D., *The Organization of Behavior*, Wiley, New York, 1949
- Hopfield, J.J., Neural Networks and Physical Systems with Emergent Collective Computational Abilities, *Proc. of the Natl. Acad. Sci.*, 1982, 79, pp. 2554-2558.
- Hoptruff, R.G.; Bramson, M.J.; Hall, T.J., Forecasting Economic Turning Points with Neural Nets, *IJCNN-91, Int. Joint Conf. On Neural Networks*, 1991, Vol.2, pp. II 1-6, Seattle, WA.
- Horng, Jyh-Ming.; Ko, Jing-Wei; I, Yet-Pole; Wu, Weng-Teng, Determination of Operating Conditions for a Crude Tower Using Neural Networks, *Process Control and Quality*, 1995, 7, pp. 21-26.
- Hornik, K.; Stinchcombe, M.; White, H., Multi-Layered Feed-Forward Neural Networks Are Universal Approximations, *Neural Networks*, 1989, 2, pp. 359-366.
- Hoskins, J.C.; Himmelblau, D.M., Artificial Neural Network Models of Knowledge Representation in Chemical Engineering, *Computers & Chem. Engng*, 1988, 12, pp. 881-890.
- Humpert, B., A Comparative Study of Neural Network Architectures, *Computer Phys. Comm.*, 1990, 58, pp.223-256
- Hunt, K.J.; Sbarbaro, D.; Zbikowski, R.; Gawthrop, P.J., Neural Networks for Control Systems – A Survey, *Automatica*, 1992, 28, pp. 1083-1112.
- Hussain, M.A.; Review of The Applications of Neural Networks in Chemical Process Control – Simulation and on-line Implementation, *Artificial Intell. In Engng.*, 1999, 13, pp. 55-68.
- Jones, W.P.; Hoskins, J., Backpropagation. A Generalized Delta Learning Rule, *Byte*, 1987, Oct., pp. 155-162.
- Kaminski, W.; Strumillo, P.; Tomczak, E., Neurocomputing Approaches to Modelling of Drying Process Dynamics, *Drying Technology*, 1998, 16, pp. 967-992.
- Karajala, T.W.; Himmelblau, D.M.; Miikkulainen, R., Data Rectification Using Recurrent (Elman) Neural Networks, *Proc. Int. Joint Conf. On Neural Networks*, 1992.

- Karjala, T.W.; Himmelblau, D.M., Dynamic Rectification of Data via Recurrent Neural Nets and the Extended Kalman Filter, *AIChE Journal*, 1996, 42, pp. 2225-2239.
- Kavuri, S.N., V., Representing Bounded Fault Classes Using Neural Networks with Ellipsoidal Activation Function, *Comp. & Chem. Engng.*, 1993, 17, pp. 139-163.
- Kennedy, M.P.; Chua, L.º, Neural Networks for Nonlinear Programming, *IEEE Trans Circuits Syst.*, 1988, CAS-35, pp. 554-562.
- Kohonen, T., *Self-Organization and Associative Memory*, 1984, Springer-Verlag, Berlin.
- Korn, G.A., Interactive Simulation of Backpropagation and Creeping-Random Search Learning in Neural Networks, *Simulation*, 1990, Oct., pp. 214-219
- Kóvacs, Z.L., *Redes Neurais Artificiais: Fundamentos e Aplicações*, 2ª ed., Ed. Collegium Cognito e Edição Acadêmica, 1996
- Kramer, M.A.; Leonard, J.A., Diagnosis Using Backpropagation Neural Network – Analysis and Criticism, *Computer & Chem. Engng*, 1990, 14, pp. 1323-1338.
- Lambert, J.; Hecht-Nielsen, R., Application of Feedforward and Recurrent Neural Networks to Chemical Plant Predictive Modeling. *Proc. Int. Joint Conf. On Neural Networks*, 1991, v. 1, pp.373-379.
- Lant, P.A.; Willis, M.J.; Montague, G.A; Tham, M.T.; Morris, A.J., A Comparison of Adaptive Estimation with Neural Based Techniques for Bioprocess Application, ACC San Diego, 1990, pp. 2173-2178.
- Latrille, E.; Corrieu, G.; Thibault, J., Neural Network Models for Final Process Time Determination in Fermented Milk Production, *Computers & Chem. Engng.*, 1994, 18, pp. 1171-1181.
- Le Cun, Y., Une Procédure d'Apprentissage pour réseau à Seuil Assymétrique, *Proc. Cognitiva*, 1985, June, pp. 599-604
- Lee, M.; Park, S., A New Control Schemes Combining Neural Network Based Feedforward Control with Model Predictive Control, *AIChE Annual Meeting*, 1991, Chicago.
- Leib, T.M.; Mills, P.L.; Lerou, J.J., Fast Response Distributed Parameter Fluidized Bed Reactor Model for Propylene Partial Oxidation Using Feed-forward Neural Network Methods, *Chem. Engng. Sci.*, 1996, 51, pp. 2189-2198.

- Lennox, B.; Rutherford, P.; Montague, G.A.; Haughin, C., Case Study Investigating the Application of Neural Networks for Process Modelling and Condition Monitoring, *Computers Chem. Engng.*, 1998, 22, pp. 1573-1579.
- Leonard, J.A.; Kramer, M.A.; Ungar, L.H., Neural Network Architecture that Computes its own Reliability, *Computers & Chem. Engng.*, 1992 a, 16, pp. 819-835.
- Leonard, J.A.; Kramer, M.A.; Ungar, L.H., Using Radial Basis Functions to Approximate a Function and its Error Bounds. *IEEE Trans Neural Networks*, 1992 b, 3, pp. 624-627.
- Lippman, R.P., An Introduction to Computing with Neural Nets, *IEEE ASSP Mag.*, 1987, Apr., pp. 4-22.
- MacMurray, J.C.; Himmelblau, D.M., Modeling and Control of a Packed Distillation Column Using Artificial Neural Networks, *Computers & Chem. Engng.*, 1995, 19, pp. 1077-1088.
- Mah, R.S.H.; Chakravarthy, V., Pattern Recognition Using Artificial Neural Networks, *Comp. & Chem. Engng.*, 1992, 16, pp. 371-377.
- Marsahll, S.J.; Harrison, R.F.; Kennedy, R., Neural Classification of Chest Pain Symptoms: A Comparative Study, *IEE Second Int. Conf. on Artificial Neural Networks*, 1991, Bournemouth, UK, pp. 200-204.
- Mavrovouniotis, M.L. & Chang, S., "Hierarchical Neural Networks", *Comp & Chem. Engng.*, 1992, 16, pp. 347-369.
- McCulloch, W. & Pitts, W., "A Logical Calculus of the Ideas Immanent in Nervous Activity", *Bulletin of Mathematical Biophysics*, 1943, 5, pp.115-133.
- Migliavacca, S.C.P., Nascimento, C.A.O. & Rodrigues, C., "Use of Neural Network for the Simulation of a Gas Centrifuge", *J. Nuclear Science. & Technology*, 1999, 36(4), 364-370.
- Miller, W.T.; Sutton, R.S.; Werbos, P.J., *Neural Networks for Control*, MIT Press, Cambridge, MA, 1990
- Minsky, M.; Pappert, S., *Perceptrons*, MIT Press, Cambridge, 1969
- Montague, G. A.; Tham, M.T.; Willis, M.J.; Moms, A.J., Predictive Control of Distillation Columns Using Dynamic Neural Network, *3rd.IFAC Symposium on*

- Dynamics and Control of Chemical Reactors, Distillation Columns and Batch Process*, 1992, pp. 231-236.
- Montague, G.A.; Hofland, A.G.; Lant, P.A.; Di Massimo, C.; Saunders, A.; Tham, M.T.; Morris, A.J., Model Based Estimation and Control: Adaptive Filtering, Nonlinear Observers and Neural Networks, Proc. 3rd. Int. Symp. Control for Profit, Newcastle-upon-Tyne, 1989.
- Montague, G.A.; Willis, M.J.; Di Massimo, C.; Morris, A.J.; Tham, M.T., Dynamic Modeling of Industrial Process with Artificial Neural Networks. *Neural Network Conference*, University of Newcastle-upon-Tyne, Newcastle, England, 1990.
- Morrison, J.D., A Neural Network Model that Supports Real Time Learning of Temporal Relationships in Complex Engineering Domains, *Simulation*, 1992, 58, pp. 152-163.
- Mukesh, D., "Hate Statistics? Try Neural Networks", *Chemical Engineering*, 1997, March, pp. 96-104
- Naidu, S.R., Zafirou, E. & McAvoy, T.J., "Use of Neural Networks for Sensor Failure Detection in a Control System", *IEEE Control Syst. Mag.*, 1990, 10, pp. 49-55.
- Narendra, K.S. & Parthasarathy, K., Identification and Control of Dynamical Systems Using Neural Networks, *IEEE Trans. Neural Network*, 1990, 1, pp. 4-27.
- Nascimento, C.A.O., Programa de redes Neurais, NEURO versão 14 – NEURO14, Departamento de Engenharia Química da EPUSP, 1992.
- Nascimento, C.A.O., Oliveros, E. & Braun, A.M., Neural Network Modelling for Photochemical Processes, *Chem. Engng. Proc.*, 1994, 33, pp.319-324.
- Nascimento, C.A.O., Guardani, R. & Giulietti, M., "Use of Neural Networks in the Analysis of Particle Size Distributions by Laser Diffraction", *Powder Technology*, 1997, vol. 90, 89-94.
- Nascimento, C.A.O. & Giudici, G., "Neural Network Based Approach for Optimization Applied to an Industrial Nylon-6,6 Polymerization Process", *Computers and Chemical Engineering*, 1998, vol. 22, 595-600.
- Nascimento, C.A.O., Giudici, G. & Scherbakoff, N. (1999). "Modeling of Industrial Nylon-6,6 Polycondensation Process. Part 2: Neural Networks and Hybrid Models.", *J. Applied Polymer Science*, 72, 905-912.

- Nascimento, C.A.O., Giudici, G. & Guardani, R., "Neural Network Based Approach for Optimization of Industrial Chemical Processes", *Computers and Chemical Engineering*, 2000, vol. 24, 2303-2314.
- Nelson, M.M., Illingworth, W.T., *A Practical Guide to Neural Nets*, Addison-Wesley Publishing Company, 1991.
- Padgett, M.L.; Roppel, T.A., Neural Network and Simulation: Modeling for Applications, *Simulation*, 1992, 58, pp. 295-305.
- Palancar, M.C.; Aragon, J.M.; Torrecilia, J.S., pH-Control System Based on Artificial Neural Networks, *Ind. Engng. Chem. Res.*, 1998, 37, pp. 2729-2740.
- Parker, D.B., Learning-Logic. *Report TR-47, Center for Computational Research in Economics and Management Science*, 1985, MIT, Mass.
- Petersen, R.; Fredenslund, A.; Rasmussen, P., Artificial Neural Networks as a Predictive Tool for Vapor-Liquid Equilibrium, *Computer Chem. Engng.*, 1993, 18, pp. S63-S67.
- Pollard, J.F.; Broussard, M.R.; Garrison, D.B.; San, K.Y., Process Identification Using Neural Networks, *Computers Chem. Engng.*, 1992, 16, pp. 253-270.
- Pottmann, M.; Seborg, D. E., A Non Linear Predictive Control Strategy Based on Radial Basis Function Networks, *3rd. IFAC Symposium on Dynamics and Control of Chemical Reactors, Distillation Columns and Batch Process 3rd.* , 1992, pp. 309-314.
- Psaltis, D.; Sideris, A.; Yamamura, A., Neural Controllers, *IEEE Int. Conf. On Neural Networks*, 1987, pp. IV-551 – IV-558.
- Psaltis, D.; Sideris, A.; Yamarura, A.A., A Multilayers Neural Network Controller, *IEEE Cont. Syst. Mag*, 1988, 8, pp. 17-21.
- Psichogios, D.C.; Ungar, L.H., Direct and Indirect Model Based Control Using Artificial Neural Networks, *Ind. Engng. Chem. Res.*, 1991, 30, pp. 2564-2573.
- Rademan, J.A.M.; Moolman, D.W.; Lorenzan, L.; van Deventer, J.S.J.; Aldrich, C., Neural Net Based Knowledge Extraction from the Historical Data of an Industrial Leaching Process, *Hydrometallurgy*, 1996, 43, pp.95-116.
- Raju, G.K.; Cooney, C.L., A Task Decomposition Approach to Using Neural Network for the Interpretation of Bioprocess Data, *Proc. IFAC Artif. Intell. In Real Time Control, Delft, The Netherlands*, 1992, pp. 447-452.

- Ramchandran, S.; Rhinehart, R.R., Very Simple Structure for Neural Network Control of Distillation, *Journal of Process Control*, 1995, 5, pp.115-128.
- Rengaswamy, R.; Venkatasubramanian, V., Syntatic Pattern-Recognition Approach for Process Monitoring and Fault Diagnosis, *Engng. Applications of Artificial Intell.*, 1995, 8, pp. 35-51.
- Rosenblatt, F., The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, *Psycholog. Rev.*, 1958, 65, pp. 386-408
- Rumelhart, D.E.; Hinton, G.E.; William, R.G. (1986), Learning Representations by Back-Propagation Errors, *Nature*, 323, pp. 533-536
- Rumelhart, D.E.; McClelland, J.L., 1986, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, 1986, MIT Press, Cambridge, MA.
- Saint-Donat, J.; Bhat, N.; McAvoy, T.J., Neural Net Based Model Predictive Control, *Int. J. Contr.*, 1991, 54, pp. 1453-1468.
- Savkovic'-Stevanovic', J., A Neural Network Model for Analysis and Optimization of Process, *Computers Chem. Engng*, 1993 a, 17, pp. S411-S416.
- Savkovic'-Stevanovic', J., Neural Networks for Process Analysis and Optimization: Modeling and Applications, *Computers Chem. Engng.*, 1994, 18, pp. 1149-1155.
- Savkovic'-Stevanovic', J., Neural Networks Models for Process Analysis, *Proc. 5th Sem. And Symp. Inform. & Exp. Sys. In the Proc Ind.*, 1991, Belgrade, Yugoslavia.
- Savkovic'-Stevanovic', J.; Vico- Stevanovic', M.; Jorgacevic',S.; Tica, N.. A Artificial Neural Network for Variable Estimation and State Identification in Process Fermentation, *Proc. of Summer Computer Simulation Conf.*, 1993 b, Boston, MA.
- Schenker, B. and Agarwal, M., "Cross-Validated Structures Selection for Neural Networks", *Comp. Chem. Engng.*, 1996, 20 (2), 175-186.
- Schizas, C.N.S.; Michaelidis, S.; Pattichis, C.S.; Livesay, R.R., Artificial Neural Networks in Forecasting Minimum Temperature, *IEE Second Int. Conf. on Artificial Neural Networks*, 1991, Bournemouth, UK, pp. 112-114.
- Scott, G.M.; Shavlik, J.W., Ray, W.H., Refining PID Controllers Using Neural Networks, *Neural Comput.*, 1992, 4, pp. 736-757.

- Sharma, R.; Singhal, D.; Ghosh, R.; Dwivedi, A., Potencial Applications of Artificial Neural Networks to Thermodynamics: Vapor-Liquid Equilibrium Predictions, *Computers & Chem. Engng.*, 1999, 23, pp. 385-390.
- Shaw, A.M.; Doyle III, F.J.; Schwaber, J.S., A Dynamic Neural Network Approach to Nonlinear Process Modelling, *Computers & Chem. Engng.*, 1997, 21, pp. 371-385.
- Spencer, P.J.; Burton, B.; Chart, T.G.; Lukas, H.L.; Pelton, A.D.; Seifert, H.; Kaufman, L.; Aldinger, F.; Fernandez Guillermet, ^a; Iorish, V., Estimation of Enthalpies and Entropies of Transformation and Trends in the Periodic Chart for the Variation of Thermodynamic Properties, *Calphad: Computer Coupling of Phase Diagrams and Thermochemistry*, 1995, 19, pp. 537-553.
- Su, H.; McAvoy, T.J., Identification of Chemical Processes Using Neural Networks, *Proc. Am. Contr. Conf.*, 1991, pp. 2314-2319.
- Su, H.; McAvoy, T.J.; Werbos, P.J., Long-Term Prediction of Chemical Processes Using Recurrent Neural Networks: A Parallel Training Approach, *Ind. Engng. Chem. Res.*, 1992, 31, pp. 1338-1352.
- Ungar, L.H.; Powell, B.A.; Kamens, S.N., Adaptive Networks for Fault Diagnosis and Process Control, *Computers & Chem Engng.*, 1990, 14, pp. 572.
- Van der Walt, T.; Barnard, E.; van Deventer, J., Process Modelling with the Regression Network, *IEEE Trans. On Neural Network*, 1995, 6, pp. 78-93.
- Venkatasubramanian, V.; Vaidyanathan, R.; Yamamoto, Y., Process Fault Detection and Diagnosis Using Neural Networks – I. Steady-State Processes, *Computers Chem. Engng.*, 1990, 14, pp. 699-712.
- Wainwright, C.E.A.; Wilson, J.A.; Pulley, R.A., Combined Neural Networks First Principles Model in Quality Control of Performance Chemicals, *Proc. 1994 IchemE Research Event*, London, 1994, pp. 832-834.
- Wang, C.J.; Wu, C.H., Parallel Simulation of Neural Networks, *Simulation*, 1991, Apr., pp. 223-232.
- Widrow, B., Generalization and Information Storage in Networks of Adaline Neurons, *em Self-Organizing Systems*, 1962

- Willis, M.J., Di Massimo, C., Montague, G.A., Tham, M.T. and Morris, A.J., Artificial Neural Networks in Process Engineerings. *IEE Proceedings-D*, 1991, 138 (3), 256-266.
- Willis, M.J.; Montague, G.A.; Di Massimo, C.; Tham, M.T.; Morris, A.J., Artificial Neural Networks in Process Estimation and Control, *Automatica*, 1992, 28, pp. 1181-1187.
- Wilson, J.A.; Zorzetto, L.F.M., A Generalised Approach to Process State Estimation Using Hybrid Artificial Neural Network/Mechanistic Models, *Computers & Chem. Engng.*, 1997, 21, pp. 951-963.
- Ydstie, B.E., Forecasting and Control Using Adaptive Connectionist Networks, *Computers Chem. Engng.*, 1990, 14, pp. 583-599.
- Zhang, J.; Morris, A.J.; Martin, E.B.; Kiparissides, C., Prediction of Polymer Quality in Batch Polymerisation Reactors Using Robust Neural Networks, *Chem. Engng. Journal*, 1998, 69, pp. 135-143.
- Zhao, H.; Hao, O.J.; McAvoy, T.J., Approaches to Modeling Nutrient Dynamics: ASM@, Simplified Model and Neural Nets, *Water Science And Technology*, 1999, 39, pp. 227-234.

- **Otimização de Colunas:**

- Edgar, T.F.; Himmelblau, D.M., Optimization of Chemical Processes, 1989, McGraw Hill, New York
- Fauth, G.F.; Shinskey, F.G., Advanced Control of Distillation Columns, *Chem. Engng. Progr.*, 1975, 71, pp. 49-54.
- Martin, G.D.; Latour, P.R.; Richard, L.A., Closed-loop Optimization of Distillation Energy, *Chem. Engng. Progr.*, 1981, Sep., pp. 33-37
- Seider, W.D., Seader, J.D. & Lewin, D.R., *Process Design Principles: Synthesis, Analysis, and Evaluation*, John Wiley & Sons, Inc., 1999.
- Shinskey, F.G., Distillation Control, 1977, McGraw Hill, New York
- Srygley, J.M.; Holland, C.D., Optimum Design of Conventional and Complex Distillation Columns, *AIChE Journal*, 1965, 11, pp. 695-701

• **Azeotropia:**

Aristovich, V.Y.; Stepanova, E.I., Determination of the Existence and Composition of Multicomponent Azeotropes by Calculation from Data for Binary Systems, *Zh. Prikl. Khim.*, 1970, 43, pp. 2192

Chapman, R.G.; Goodwin, S. P., A General Algorithm for the Calculation of Azeotropes in Fluid Mixtures, *Fluid Phase Equilibria*, 1993, 85, pp. 55

Ewell, R.H.; Harrison, J.M.; Berg, L., Azeotropic Distillation, *Ind. Eng. Chem.*, 1944, 36, pp. 871-875.

Fidkowski, Z.T.; Malone, M.F.; Doherty, M.F., Computing Azeotropes in Multicomponent Mixtures, *Comput. Chem. Eng.*, 1993, 17, pp. 1141.

Gmehling, J., Böltz, R., Azeotropic Data for Binary and Ternary Systems at Moderate Pressures, *J. Chem. Eng. Data*, 1996, 41, pp. 202-209.

Gmehling, J.; Menke, J.; Krafczyk, J.; Fischer, K., A Data Bank for Azeotropic Data – Status and Applications, *Fluid Phase Equilibria*, 1995, 103, pp. 51-76.

Gmehling, J.; Menke, J.; Krafczyk, J.; Fischer, K., *Azeotropic Data*, 2 vol., CH-Verlag, Weinheim, 1994.

Guggenheim, E. A. et al., *Thermodynamics*, North-Holland, Amsterdam, 1967.

Harding, S.T.; Maranas, C.D.; McDonald, C.M.; Floudas, C.A., Locating All Homogeneous Azeotropes in Multicomponent Mixtures, *Ind. Eng. Chem. Res.*, 1997, 36, pp. 160-178.

Horsley, L.H., Azeotropic Data – III. *Advances in Chemistry Series*, No. 116. American chemical Society, Washington DC, 1973.

Kim, Y.J. & Simmrock, K.H., “AZEOPERT: An Expert System for the Prediction of Azeotrope Formation – I. Binary Azeotropes”, *Computers Chem. Engng*, 1997, 21, pp. 93-111.

Lecat, M., *Tables Azeotropiques*. Monograph, Brussels, 1949

Maier, R.W.; Brennecke, J.F.; Stadtherr, M.A., Reliable Computation of Homogeneous Azeotropes, *AIChE Journal*, 1998, 44, pp. 1745-1755.

Malesinski, W., *Bull. Acad. Polon. Sci. Cl.III*, 1956, 4, pp. 295, 303, 365, 371.

- Maranas, C.D.; McDonald, C.M.; Harding, S.T.; Floudas, C.A., Locating All Azeotropes in Homogeneous Azeotropic Systems, *Comput. Chem. Eng.*, 1996, 20, pp. S413.
- Michelsen, M.L., The Isothermal Flash Problem: Part I. Stability, Fluid Phase Equilibria, 1981, 9 pp. 1-19.
- Okasinski, M.J.; Doherty, M.F., Thermodynamic Behavior of Reactive Azeotropes, *AIChE Journal*, 1997, 43, pp. 2227.
- Swietoslawski, W., *Bull Acad. Polon. Sci.*, 1950, 19, pp. 29.
- Teja, A.S.; Rowlinson, J.S., The Prediction of the Thermodynamic Properties of Fluids and Fluid Mixtures – IV. Critical and Azeotropic States, *Chem. Eng. Sci.*, 1973, 28, pp. 529
- Trotsenko, A.V., Prediction and Calculation of Azeotropic Behaviour from an Equation of State, *Fluid Phase Equilibria*, 1997, 127, pp. 123-127.
- Wang, S.; Whiting, W.B., New Algorithm for Calculation of Azeotropes from Equations of State, *Ind. Eng. Chem. Proc. Des. Dev.*, 1986, 25, pp. 547.
- Westerberg, A.W.; Wahnschafft, O., The Synthesis of Distillation Based Separation Systems, *Adv. Chem. Eng.*, 1996, 23, pp. 64.
- Widagdo, S.; Seider, W.D., Azeotropic Distillation, *AIChE Journal*, 1996, 42, pp. 96-130.
- Yoshimoto, T.; Mashiko, Y.I., "Studies on Azeotropic Mixtures I", *Bull. Chem. Soc. Jpn.*, 1956, 29, pp. 990.

- **Diversos:**

- Abraham, M.H., *Chem. Soc. Rev.*, 1993, 22, 73.
- Abraham, M.H., Chadham, H.S., Whiting, G.S. & Michell, R.C., *J. Pharm. Sci.*, 1994, 83, 1085.
- Guardani, R., *Programa para separação dos dados em conjunto de Aprendizado e Teste, PREP 2. v.1*, Departamento de Engenharia Química da Escola Politécnica da Universidade de São Paulo, 1998.

- Kachigan, S.K., *Multivariate Statistical Analysis-A Conceptual Introduction*, 2nd ed., Radius Press, New York, (1991).
- Khattree, Ravindra & Naik, D. N., *Multivariate Data Reduction and Discrimination with SAS Software*. SAS Institute Inc., North Carolina, 2000.
- Nascimento, C.A.O., *Programa de Redes Neurais, Neuro14*, v.14, Departamento de Engenharia Química da Escola Politécnica da Universidade de São Paulo, 1991.
- Nascimento, C.A.O ; Alves, R.M.B., *Programa para Simulação do Processo de Produção de Isopreno via Redes Neurais*, v.1, Departamento de Engenharia Química da Escola Politécnica da Universidade de São Paulo, 2003
- Nascimento, C.A.O; Alves, R.M.B., *Programa para Otimização do Processo de Produção de Isopreno via Redes Neurais*, v.1, Departamento de Engenharia Química da Escola Politécnica da Universidade de São Paulo, 2003
- Neto, J.P.P.; Carneiro, W.L. *Otimização da Coluna Depentanizadora para Produção de Piperilenos*, Estudo Técnico, COPENE PETROQUÍMICA DO NORDESTE S.A, 1997
- Pellicci, R.L.; Carneiro, W.L., *Otimização do Reator de Dimerização de CPD*, Estudo Técnico, COPENE PETROQUÍMICA DO NORDESTE S.A, 1996
- Reid, R.C., Prausnitz, J.M.; Poling, B.E., *The Properties of Gases & Liquids*, 1988, McGraw-Hill Int. Ed., 4th ed.

APÊNDICE A – ARTIGOS

Todos os artigos referentes a trabalhos apresentados em Congressos, publicados e submetidos para publicação em periódicos são apresentados neste apêndice.

GROSS ERROR DETECTION OF INDUSTRIAL DATA BY NEURAL NETWORK AND CLUSTERS TECHNIQUES

R. M. B. Alves* and C. A. O. Nascimento

LSCP- Laboratório de Simulação e Controle de Processos,
Departamento de Engenharia Química
Escola Politécnica da Universidade de São Paulo
Av. Prof. Luciano Gualberto, n. 380, trav. 3, 05508-900, São Paulo – SP, Brasil
E-mail: rita@lscp.pqi.ep.usp.br; oller@usp.br

Abstract. This article describes the application of a three layer feed forward neural network to industrial plant data analysis. To adjust mathematical models (for control or optimization purpose) from plant data, it is necessary to analyze and detect outliers and systematic errors and remove them. The system studied is the feed preparation of an Isoprene Production Unit and represents a multivariable problem. To detect outliers on a multivariable system is not an easy task. The technique used in this paper is able to identify this kind of error. The methodology employed involves the construction of a reliable neural network model to represent the process and its training with few iterations (few thousands times). Then, the points in which errors between the experimental and calculated data appear to be scattered far away from the majority of values are probably outliers. For some cases, outlier points can be easily detected, but for others, it is not so obvious. For these, they are separated and a cluster with others similar data are built. Analyzing these clusters with basis on the similarity principle or by hypothesis tests for means, it is then decided if these points can be excluded or not. In parallel, it is verified if any abnormality on the process was recorded on the specific period. Three years of process data were treated and about 30% of the data were excluded.

Keywords: Gross Error, Neural Network, Modeling, Data Analysis

1. Introduction

Gross errors or anomalous measurements to data set may arise due to changed conditions during a plant operation, or due to errors at the operation of measurements and recording devices, or simply due to error on the information register, which may contaminate the valid data. According to the average time considered on the data treatment, data fluctuation could be incorporated on the results. Many times, this could lead to unreliable information. In cases of errors of the measurement instruments at a long period of time, the average reflects this error. On the other hand, the outlier may be simply one of the extreme values in a probability distribution for a random variable, which occurs quite naturally but not frequently and should not be rejected.

* To whom all correspondence should be addressed

If it knows the origin of the abnormal values, it does not hesitate to discard such observation. On the other hand, when it is not sure about the error or it does not have enough practice to either accept or reject an extreme observation, he must based his judgment on some kind of statistical analysis. The question to be analyzed is how probable it is that the observed differences are due solely to random sampling errors in order to reject or not the information. This task becomes especially complicated for complex processes where not all of the influencing parameters are directly accessible or where large stochastic deviations of the process variables lead to a considerable scatter of the measured data. For this reason, a large variety of approaches were proposed in the past, which tackle this problem. These are commonly based on either statistics or first principle equations or a composition of both. However, this procedure may become extremely complicated both, if the underlying physics and chemistry of the process are not very well understood, and if the application of a sharp statistical criterion for the separation of the data into one set of valid and another of non-valid values is impossible.

This article demonstrates the ability of a neural network model in learning and adapting themselves to different statistical distributions of inputs involving non-linear mappings. In this way, it allows the classification of similar inputs and outputs in order to identify clusters and then proceed with the elimination of the gross errors. As will be shown, this approach to detect outliers possesses a considerable potential in the field of data analysis, it is easier and requires much less knowledge of the underlying physical-chemical process.

2. Neural Network

Neural computation has become an established discipline and has gained extensive interest within chemical engineering. Most chemical engineering processes are non-linear and complex with conventional modeling and simulation techniques relying often on certain simplifying transport, kinetic and/or thermodynamic assumptions. Artificial neural networks (NN), on the other hand, are able to extract information from data plant in an efficient manner. NNs have been successfully employed in solving problems in areas such as fault diagnosis, dynamic modeling and control of chemical processes (Bhat and McAvoy, 1990; Hoskins and Himmelblau, 1988; Giudici et al., 1999), and to solve nonlinear optimization problems (Nascimento & Giudici, 1998, Nascimento et al., 2000), among others. In spite of the extensive range of NN applications explored for use in chemical engineering, the quality of information is crucial for the net training and also to avoid overfitting.

Artificial neural networks are made up of highly interconnected layers of simple neuron-like nodes. The neurons act as non-linear processing elements within the network. An attractive property of artificial neural networks is that, given the appropriate network topology, they are capable of characterizing non linear functional relationships, representing internal models of a system through a direct learning algorithm, and so they are able to handle the intrinsic complexities of chemical processes. Among the many existing artificial neural network paradigms, the three layers feed forward neural network is the most widely used network for chemical engineering applications. This NN is classified as supervised learning network, in which knowledge is captured in the strength of its interconnections between a set of artificial neurons. These interconnections are called weights of neural model, which are calculated iteratively using a backpropagation algorithm, i.e. steepest descent-based optimization routine in order to minimize a given objective function (Rummelhart & McClelland, 1986). The computations are carried out

over the entire network, except the input layer. The mapping of each unit is in terms of the combination of all its inputs, followed by the application of a non-linear function, called activation function. In this work, a sigmoid function was used as activation function.

The construction and training of the NN, used in this work, were carried out using an in-house software.

3. Methodology

The available industrial data of the studied process (Isoprene Production Unit) were provided as daily average. These data were collected during three years. The analysis was carried out for each year individually. The treatment of the data was performed in two steps: first, a preliminary analysis for abnormal values was carried out, e.g. points out of the possible operation range, which may be subject to rejection. The second step involves the data analysis by using neural network approach and statistical techniques. The methodology applied at this work follows the structure shown in Figure 1 below:

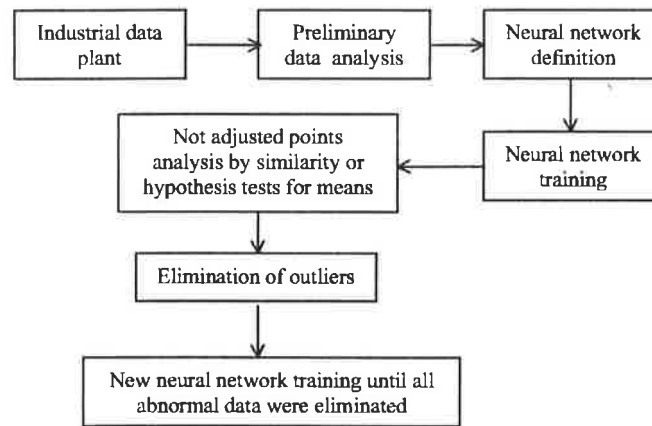


Fig.1. Data analysis methodology

The first step in the data analysis makes use of the following criteria to eliminate the abnormal values: values out of an acceptable range for the corresponding variable; graphic observations of the variables as time function; experience whether statistical features as the process; material and energy balances. The variables of interest were defined considering the available process data and its importance for the process and plant operation. Then, the minimum, maximum and mean values were identified as well the variance for each selected variable. The variables whose operational range are too close to the instruments limits (e.g. wind-up measurements) were not included as neural network information.

After this previous analysis, the resulting data set was used to train the neural network, as shown in Figure 1. At this step, all data were included in the training data set and during the training process few iterations were performed. It was observed that some points and groups of points were not well adjusted. The identification of these groups (or points) is an indication of consistency problems or gross errors not identified at first step procedure. To

decide if these points must or must not be eliminated it was make use of some statistical analyses as clusters analysis and hypothesis tests for means.

The cluster analysis is based on the similarity principle among several data sets. A data set was formed by the input and output variables chosen for each process unit, corresponding to the one operation day information. It is expected that for a series of similar input variables, the process must yield similar output variables (dependent variables). When it is observed a different input or output variable among a series of similar data, the corresponding data set may be rejected. Table 1 shows two examples of clusters analysis: it can be observed, for the variable out2, that the value 25.24 and 20.85, in the first and second group of data respectively, must be rejected.

Table 1. Clusters Analysis - Examples

		Input Variables						Output Variables				
	in1	in2	in3	in4	in5	in6	in7	in8	out1	out2	out3	out4
	12.79	15.24	1.86	63.80	39.20	59.90	8.87	4.35	8.35	25.24	2.63	97.3
	12.80	15.40	1.88	63.80	39.20	59.90	8.85	4.38	8.35	26.88	2.53	97.4
	12.75	15.07	1.71	63.80	39.10	59.80	8.79	4.36	8.35	27.21	2.49	97.3
	13.99	16.03	2.40	64.50	55.10	59.80	11.19	4.83	9.08	20.85	2.68	102.6
	13.82	16.15	2.93	64.30	55.50	59.30	10.79	5.06	9.16	28.84	2.89	97.2
	13.80	16.09	2.91	64.30	55.10	59.40	10.81	4.97	9.23	28.57	2.85	97.8
	13.70	16.02	2.74	64.20	55.60	58.60	11.13	4.91	9.07	30.50	2.66	95.0
	13.70	15.86	2.54	64.20	55.80	58.60	11.15	5.07	8.92	30.92	2.54	94.6

Table 2. Hypothesis Test for Means

Group	Input Variables							Output Variables				
	in1	in2	in3	in4	in5	in6	in7	in8	out1	out2	out3	out4
1	13.00	14.50	1.95	66.00	51.10	59.90	10.51	5.55	7.67	26.73	1.62	93.5
	13.00	14.57	2.77	65.90	50.60	60.00	10.25	5.37	7.72	26.15	2.68	91.6
2	13.20	16.20	2.43	64.30	50.40	59.70	9.94	4.41	8.68	25.81	3.17	100.1
Sample	12.00	13.58	2.38	66.00	50.20	61.20	10.40	3.77	8.78	19.91	2.77	97.2
	12.00	14.74	2.14	65.90	48.40	61.30	10.08	3.77	8.89	20.77	2.70	98.5
	12.09	14.13	1.75	66.50	48.60	61.00	10.23	4.04	8.58	23.51	2.07	97.5
	13.30	14.99	2.46	66.80	48.80	60.90	10.32	5.15	8.73	24.35	2.70	94.8
	13.48	15.74	2.37	66.80	49.60	60.70	10.47	5.35	8.76	24.60	2.47	93.8
	13.50	15.64	2.23	66.80	50.30	60.80	10.57	5.48	8.72	25.22	2.30	93.8
	13.51	15.02	2.01	66.70	50.50	60.80	10.56	5.51	8.76	25.23	2.26	94.0
	13.50	15.87	1.96	66.40	49.10	60.60	10.34	5.49	8.79	25.91	2.17	94.5
	13.00	14.50	1.95	66.00	51.10	59.90	10.51	5.55	7.62	26.73	1.62	93.5
	12.98	13.52	2.22	65.70	51.60	60.00	10.65	5.38	7.74	26.09	2.57	93.1
	13.00	14.57	2.77	65.90	50.60	60.00	10.25	5.37	7.72	26.15	2.68	91.6
	13.20	16.20	2.43	64.30	50.40	59.70	9.94	4.41	8.68	25.81	3.17	100.1
	13.00	14.06	2.05	64.30	51.20	59.70	10.27	4.28	8.65	20.93	2.66	100.5
	13.38	14.13	1.98	64.30	50.10	59.40	10.38	4.41	8.84	20.98	2.52	100.8
	14.01	13.99	2.60	64.50	52.30	59.60	10.97	4.40	9.45	17.98	2.94	100.9
	14.00	14.24	2.64	64.40	52.30	59.50	11.00	4.61	9.30	18.49	2.89	99.6
	14.00	14.11	2.57	64.40	52.30	59.50	11.00	4.62	9.30	19.09	2.89	99.9
	14.01	15.67	2.59	64.30	50.70	59.70	10.81	4.56	9.41	19.65	2.82	99.8
Minimum	12.00	13.52	1.75	64.30	48.40	59.40	9.94	3.77	7.62	17.98	1.62	91.60
Maximum	14.01	16.20	2.77	66.80	52.30	61.30	11.00	5.55	9.45	26.73	3.17	100.90
mean	13.22	14.71	2.28	65.56	50.45	60.24	10.49	4.79	8.71	22.86	2.57	96.88
Std dev	0.65	0.82	0.29	1.04	1.23	0.66	0.31	0.63	0.54	3.06	0.37	3.20
t/n-1=18	2.11											
Mean+std dev*t	14.60	16.44	2.90	67.75	53.05	61.63	11.13	6.10	9.84	29.31	3.35	103.63
Mean-std dev*t	11.84	12.97	1.67	63.37	47.85	58.85	9.84	3.47	7.57	16.41	1.78	90.13

In some cases a simple and direct analysis is not possible, e.g. when a given data set is unique or when there are only two data set for comparison with some distinct information, then it is not possible to determine which one is correct. In these cases, the domains of the variables are extended compared with the previous group. Although these new groups are less accurate, usually it is possible to discriminate the abnormal point. For this step, the hypothesis test for means was employed, which involves the confidence interval estimate and hypothesis test, with 95% as confidence level (Himmelblau, 1970).

Table 2 shows the application of this methodology on the data plant analysis. The values in bold in groups 1 and 2 were not well adjusted during the neural network training and it was not possible identify groups of similar data set for the clusters analysis, then, the hypothesis test for means analysis was performed. Its was observed that the value 25.81 is inside the interval of confidence and the null hypothesis is accepted and this data must not be eliminated. On the other hand, the value 1.62 is outside the confidence interval and the null hypothesis is rejected and this data set must be eliminated.

4. Results and Discussion

A simple measure to assess the quality of fitness of the chosen neural network to the experimental data is usually a comparison of the values calculated by the neural network to the original experimental data. The scatter of data points around the ideal 45° -line can be used to judge the fit of the neural network to the experimental data. The idea to use neural networks for the purpose of outlier detection is based on this kind of diagram (Büllau et al., 1999). Hence, it only has to be shown that the outliers of the experimental data correspond to the outliers from this curve. Thus, the neural network was first of all trained for the entire data set and afterwards for the filtered data set. The outliers detected after the first training run were analyzed by the application of the statistics described before. This procedure was repeated several times until the scattered data did not show abnormal points. Since the training of the network with the filtered data leads to different results for the calculated data in comparison to the original data set, the input data base changes due to the filtration procedure. Table 3 shows the number of eliminated points at each run and Figures 2a-c show the results of this method for the first, second and final runs, respectively.

Table 3. Eliminated points at each run

Run	Number of eliminated points
1	49
2	28
3	19
4	7
5	12
6	-

5. Conclusions

The application of a neural network is a very attractive tool to detect outliers. It is simple, more cost effective and may be easier to be used, particularly by plant engineers and the results presented demonstrate that neural networks posses a considerable potential in the field of data analysis, mainly because it requires much less knowledge of the

underlying physical-chemical process. However, the final decision to eliminate the suspect data is made by applying the cluster analysis this approach to detect outliers.

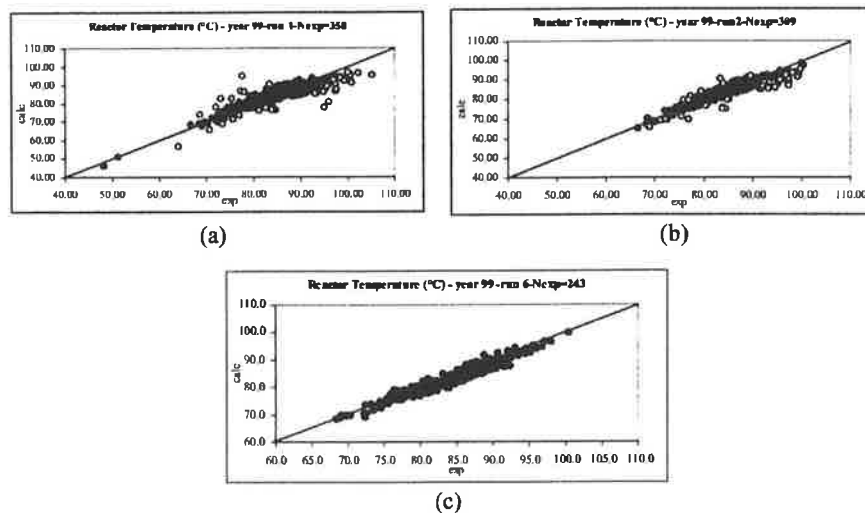


Fig. 2. Comparison of calculated and measured data: (a) before analysis; (b) intermediate results; (c) final result

Acknowledgment

The authors gratefully acknowledge FAPESP for their financial support and COPENE Petroquímica do Nordeste for providing the industrial data used in this work.

References

- Bhat, N., McAvoy, T. J. (1990). Use of Neural Nets for Dynamic Modeling and Control of Chemical Process Systems. *Comput. Chem. Eng.*, 14 (4/5), 575.
- Bülau, H. C., Ulrich, J., Guardani, R., Nascimento, C. A. O. (1999). Application of Neural Networks to Data from a Melt Crystallization Process for the Detection of Outliers. In Proceedings of AIDA, International Seminar on Advances in Data Analysis – Washington, DC
- Giudici, R., Nascimento, C. A. O., Tresmondi, A., Domingues, A., Pellicciotta, R. (1999). Mathematical Modeling of an Industrial Process of Nylon-6,6 Polymerization in a Two-Phase Flow Tubular Reactor. *Chem. Eng. Sci.*, 54, 3242.
- Himmelblau, D. M. (1970). *Process Analysis by Statistical Methods*. John Wiley & Sons, Inc.
- Hoskins, J. C., Himmelblau, D. M. (1988). Artificial Neural Network Models for Knowledge Representation in Chemical Engineering. *Comput. Chem. Eng.*, 12 (9/10), 881.
- Nascimento, C. A. O., Giudici, R. (1998) Neural Network Based Approach for Optimization Applied to an Industrial Nylon-6,6 Polymerization Process. *Comput. Chem. Eng.*, 22, S595.
- Nascimento, C. A. O., Giudici, R., Guardani, R. (2000). Neural Network Based Approach for Optimization of Industrial Chemical Processes. *Comput. Chem. Eng.*, 24, 2303.
- Rummelhart, D. E., McClelland, J. L. (1986). *Parallel Distributed Processing* (Vol. 1), Cambridge, MA: MIT Press

Neural Network Based Approach for an Industrial Isoprene Process Optimization

R. M. B. Alves and C. A. O. Nascimento

LSCP- Processes Control and Simulation Laboratory,
Department of Chemical Engineering
Polytechnic School, University of São Paulo
Av. Prof. Luciano Gualberto, n. 380, trav. 3, 05508-900, São Paulo – SP, Brazil
E-mail: rita@lscp.pqi.ep.usp.br; oller@usp.br



This article describes the application of a neural network (NN) to an industrial plant for isoprene production. The purpose of this plant is to produce high purity isoprene for obtaining synthetic rubber and thermoplastic rubber from a C_5 cut arising from a pyrolysis gasoline unit. The process of isoprene production can be divided into five sections: feed preparation, extractive distillation, and solvent recovery, sulfur removal and fractionating. The system consists basically of a dimerization reactor and a separation column train. The basic idea is the optimization of this chemical process using NN, which replaces phenomenological models and following a grid search in the region of interest [1]. Since most chemical processes are non-linear, NN appears as a technique of interest due to its capability of learning the system without knowledge of the physical and chemical laws that govern it. Moreover, the NN readily deals with constraints, avoids several typical numerical problems of conventional optimization tools, and is not computationally time-consuming. However, success in obtaining a reliable and robust NN depends strongly on choosing of the process variables involved, as well as the quality of available data set and the domain used for training purposes. For this work the neural network model is built directly from the historical plant data, which were collected during a period of three years. These data were carefully analyzed in order to identify and eliminate data containing gross errors or those on non-steady state operation. Comparisons between the model's prediction and these data were performed and reasonable results were achieved from industrial point of view. This new approach provides comprehensive information for an engineer's analysis and the use of NN methodology is promising for industrial applications by changing variable constraints.

INTRODUCTION

Methods for optimization of a given process involve the measure of goodness, or objective function, equality and inequality constraints, as well as constraints related to safety considerations and those that arise from the process model equations. The objective function is established from a technical and/or economic viewpoint. The equality constraints usually represent specifications while the inequality constraints usually represent the lower and upper bounds of the operating variables.

Formal methods of optimization can be utilized to optimize a superstructure of process units with streams that can be turned on and off using binary variables. The mixed-integer formulation of the optimization problem, in principle, permits the optimizer to select simultaneously the best flowsheet and optimize it with respect to its continuous variables, such as the pressure levels, reflux ratios, residence times, and split fractions. In practice, however, most design problems are not solved using superstructures and mixed-integer optimization algorithms. Rather, heuristics together with simulation and algorithmic methods are utilized to build and analyze synthesis trees. Although substructures, such as networks of heat exchangers, can be optimized conveniently using mixed-integer methods, it is impractical to attempt the optimization of entire process flowsheets in this manner [2].

Process simulators have been used to carry out the optimization while converging the recycle loops. In this case, the largest fraction of equality constraints are modeling equations associated with the process units. For example, a distillation column that is modeled with stages assumed to be in phase equilibrium often has several hundred MESH equations. However, in the implementation of most simulators, these equations are solved for each process unit, given equipment parameters and stream variables (usually for feed streams), using subroutines in a program library. Hence, when using these simulators, the equations (or equality constraints) for the process units are shown explicitly in a nonlinear program. Given values for the design variables, the simulators call upon these subroutines to solve the appropriate equations and obtain the unknowns that are needed to perform the optimization. This method is time-consuming since it usually requires a higher number

of iterations, Moreover, in process design and optimization, up to 80% of the total computational time can be spent on evaluation of thermodynamic properties [2].

Neural networks (NN) appear as one of the most promising techniques to model and simulate industrial chemical processes [3]. NN replace the phenomenological models due their capability of learning the system without knowledge of the physical and chemical laws that govern it. Moreover, the NN readily deal with constraints, avoid several typical numerical problems of conventional optimization tools, and are not computationally time-consuming. However, the success in obtaining a reliable and robust NN depends strongly on choosing of the process variables involved, as well as the quality of the available data set and the domain used for training purposes.

PROCESS DESCRIPTION

The objective of the Isoprene Extraction Unit is to yield high purity isoprene from a C₅ cut of pyrolysis gasoline. The isoprene produced is to be used to obtain synthetic and thermoplastic rubber. The process of isoprene production can be divided into five sections: feed preparation, extractive distillation and solvent recovery, sulfur removal and fractionating. The system consists basically of a dimerization reactor and a separation column train. This work deals with the optimization of the feed preparation section. At this section, the unit feed stream flows through a tubular reactor to thermally dimerize cyclopentadiene (CPD) to dicyclopentadiene (DCPD). The reactor effluent is fed to a distillate column for removal of heavies. This first column separates one stream rich in isoprene, as distillate, from a heavy cut, rich in DPCD and piperylenes, as bottom product. This bottom stream feed another column to separate the heavy cut in two streams: one rich in DCPD and other rich in piperylenes, both used as raw material in resin manufacture. The distillate product of the first column is the feed stream of the isoprene extraction section, which uses a mixed solvent to separate high purity isoprene from the other compounds of the C₅ stream. Figure 1 shows schematically the feed preparation section.

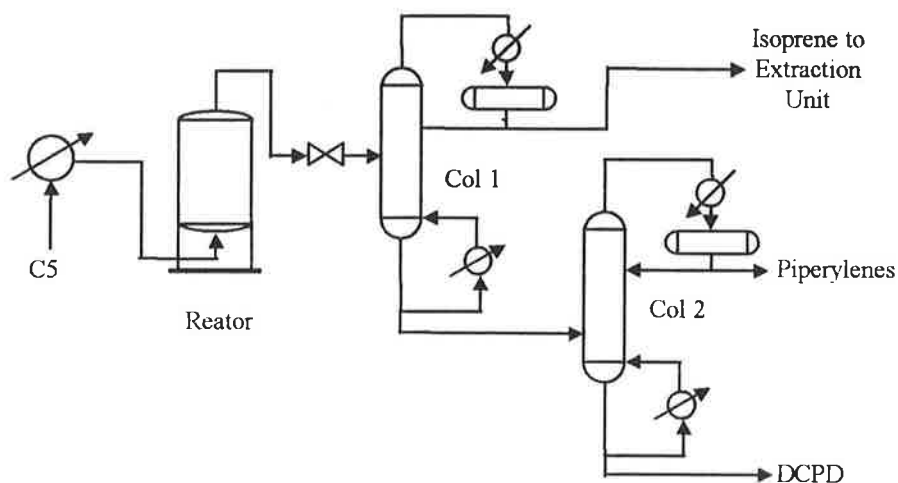


Figure 1. Feed Preparation Flow Diagram

NEURAL NETWORK MODELS

Neural networks have become a widely used tool in areas such as process analysis, simulation, and control, primarily due to their inherent advantages such as e.g. adaptability, non-linearity, and fault tolerance specially in combination with the relatively simple way of use.

The neural network employed is the feed forward network, in which information propagates in only one direction and is useful for steady-state modeling. The architecture of a typical NN is shown in Figure 2. The nodes in different layers of the network represent "neuron-like" processing elements. The number of neurons in input and output layers depends on the respective number of inputs and outputs being considered. On the other hand, hidden layers may vary, depending on the user specification, and its structure defines the feed forward

NN topology. The neurons in the input layer do not perform data processing functions; they merely provide a means by which scaled data is introduced into the network. These signals are then “fed forward” through the net via the connections, through hidden layers, and eventually to the final output layer. Each interconnection has associated with it a weight that modifies the strength of the signal flowing along that path. Thus, the input to each neuron is a weighted sum of the outputs from neurons in the previous layer. The output of each node is obtained by passing the weighted sum through a nonlinear operator, in this work, a sigmoidal function. This function provides the network with the ability to represent a nonlinear relationship. Selection of an optimal network structure is crucial for network performance. For example, a network with too few hidden nodes can not represent the system satisfactorily and a network with too many hidden nodes overlearns particular runs presented in training set.

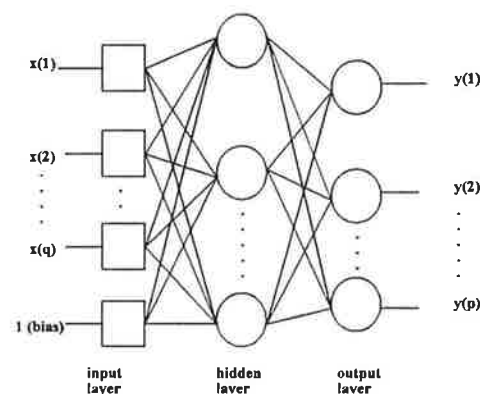


Figure 2. Structure of a feed forward neural network

Once the network topology is specified, a set of input-output data is used to train the network; i.e. to determine appropriate values for the weights associated with each interconnection. The data are propagated forward through the network to produce an output that is compared with the corresponding output in data set, hence generating an error. This error is minimized by making changes to the weights, and may involve many passes through training data set. For a given topology, the magnitude of the weights defines the network characteristics and the structural properties of the model. Thus, a NN has the capability to represent complex systems whose structural properties are unknown [4].

The typical way of network design consists of training various network structures with one part of the data, the so-called training set, and their performance is evaluated afterwards with the remaining part of the data, the so-called test set. Both sets can be made large enough to ensure satisfactory learning of the network weights and accurate evaluation of the network performance [5].

NEURAL NETWORK BASED APPROACH FOR OPTIMIZATION

The main idea is to replace the model equations with an equivalent neural network, and use this NN to carry on a detailed grid search of the region of interest, by a full mapping of the objective function on the space of decision variables. Using NN instead of a phenomenological model itself takes advantage of the comparative fast response by a neural network simulation. In this way, even a detailed grid search can be achieved in reasonable time, as long as there are not too many variables being optimized (in which case problems of dimensionality arise). This approach is more reliable, readily deals with constraints, avoids several typical numerical problems of conventional optimization tools and is not computationally time-consuming. As an additional benefit, full mapping of the objective function allows one to identify multiple optima easily, an important feature not presented by conventional optimization methods. Moreover, the constraints are easily treated afterwards since points with violated constraints can be recognized and classified (according to weak or hard constraints) [1].

Once the map is obtained, it is easy to choose the optimum point, to identify whether multiple optima are present, to check if constraints were violated, and so on. This approach definitely provides more comprehensive information for an engineer's analysis than the conventional non-linear programming

procedure. This algorithm can be straightforwardly extended to treat a multi-objective optimization problem as well.

METHODOLOGY

The structure of the modeling using the neural network approach is shown in Figure 3, and the global structure of optimization is shown in Figure 4.

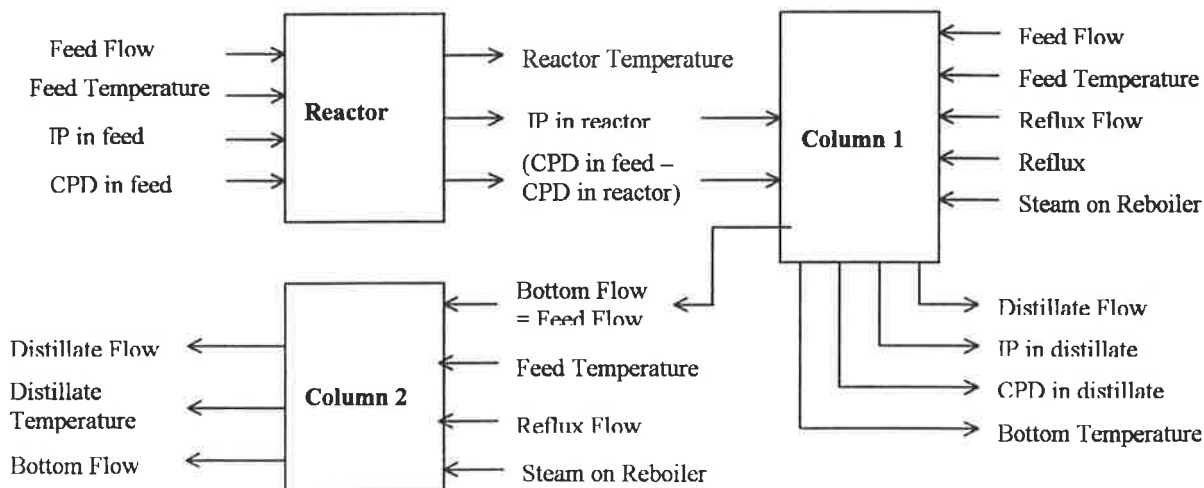


Figure 3. Scheme of the NN modeling

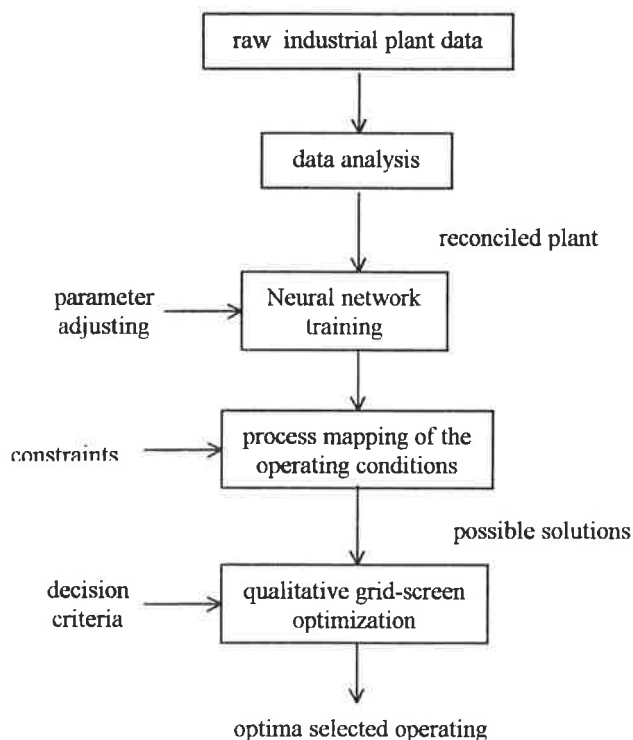


Figure 4. Optimization procedure scheme

Data Preparation

The neural network modeling was done directly from industrial data. Daily averages of each variable were provided, for a period of three years of plant operation. However presence of noise and gross errors,

commonly found in industrial processes would disturb the modeling and lower its quality. Therefore, for this work, the first step was to prepare the industrial data to a model fitting by detection and elimination of data containing gross errors ("outliers"). This was done with basis at cluster analysis by similarity or hypothesis test for means as shown in [6]. By applying such procedures, about 30% of the plant data were eliminated.

A large number of process variables were available to fit the neural network model. Table 1 shows the variables selected and its applicability range:

Table 1. Process variables and its application range

Process Unit	Variables	Notation	Unit	Maximum	Minimum
Reactor	Feed flow	x1(1)	t/h	14,51	9,49
	Feed temperature	x1(2)	°C	79,0	62,0
	IP in feed	x1(3)	%weight	23,68	13,48
	CPD in feed	x1(4)	% weight		
	Reactor temperature	y1(1)	°C	100,5	68,3
	IP in reactor	y1(2)	%weight	21,64	12,27
	CPD feed-CPD reactor	y1(3)	%weight	13,59	21,64
Column 1	Feed flow	x2(1)	t/h	14,51	9,50
	Feed temperature	x2(2)	°C	66,5	62,6
	IP in feed	x2(3)	%weight	21,48	12,27
	CPD in feed	x2(4)	%weight	4,74	2,01
	Reflux flow	x2(5)	t/h	66,3	35,1
	Reflux temperature	x2(6)	°C	61,20	56,9
	Reboiler Steam	x2(7)	t/h	11,90	8,41
	Distillate flow	y2(1)	t/h	9,10	5,92
	IP in distillate	y2(2)	%weight	38,72	23,55
	CPD in distillate	y2(3)	%weight	4,43	1,87
	Bottom temperature	y2(4)	°C	103,8	88,6
	Bottom flow	y2(5)	t/h	6,30	2,75
Column 2	Feed flow	x3(1)	t/h	6,58	3,14
	Feed temperature	x3(2)	°C	61,9	52,9
	Reflux flow	x3(3)	t/h	4,50	3,99
	Reboiler Steam	x3(4)	t/h	1,45	1,12
	Distillate flow	y3(1)	t/h	2,49	1,55
	Bottom flow	y3(2)	t/h	4,28	1,58
	Distillate temperature	y3(3)	°C	52,6	48,4

Modeling

After the plant data analysis, the process modeling was performed. This step consisted of establishing an optimal neural network structure to represent each process unit individually.

Initially, the industrial data was divided into two sets: a learning set and a test set. The test set contained variables within the range of the learning set variables. Thus, no extrapolation takes place in the model by comparison with the test set data.

Then, the NN was run with several different specifications, i.e., varying the number of neurons in the hidden layer and the number of data presentations (iterations), in order to select the best one. To avoid an overfitting of the model, choosing a minimum neuron number in the hidden layer (NH), as well as an adequate number of training cycles that accomplishes the lowest error in the test set is recommended [7]. Figure 5 shows an example of this procedure, i. e., the minimum error for the test set as a function of the number of neurons and presentations. A minimum value for error is reached for NH equals 9 and about 30000 presentations can be observed.

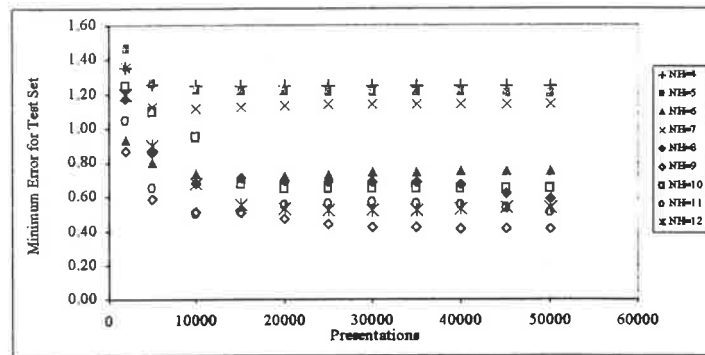


Figure 5. Minimum error for the test set

For the present work, the selected neural network configurations are shown in Table 2 for each process unit involved in this study.

Table 2. Neural network configurations

	Neurons in hidden layer	Presentations	Number of data set in learning set	Number of data set in test set
Reactor	9	30000	165	78
Column1	12	30000	156	63
Column2	13	45000	164	75

The next step was to verify the neural network model by comparing measured and calculated output variable values. Figures 6a and 6b show some examples of this comparison for the test data set. Good agreement between modeled and measured values can be observed and the uniform distribution indicates absence of any tendency on fitting.

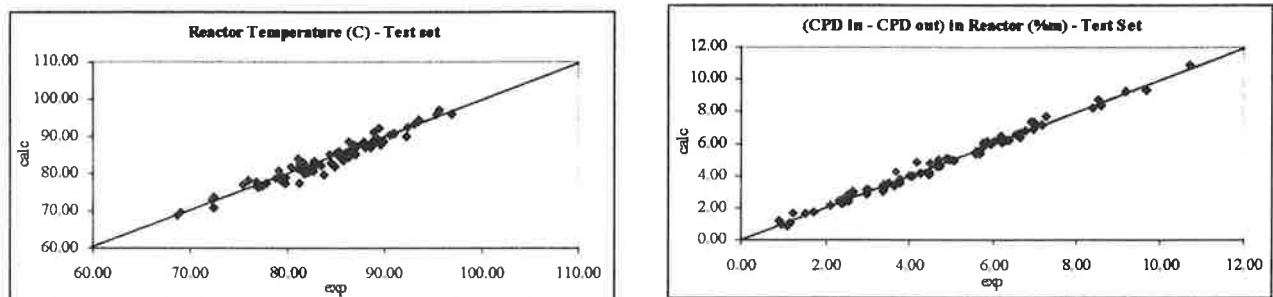


Figure 6. Comparison between experimental and calculated values

Optimization

The first step is to give the feed specification to the reactor, i. e., total flow rate and feed concentration of isoprene (IP) and CPD. Then the mapping of solution is initiated by changing the process variables and at this step only the solutions which satisfy some operational conditions are accept. These conditions are CPD outlet concentration and IP loss from the reactor that must be in a given range. IP and CPD concentrations in the overhead product from the distillation column must be in a given range also.

The next step is choose qualitatively the operational conditions which could minimize the energy costs, these conditions are the lowest reflux flow, the lowest reboiler steam flow and the lowest feed temperature. Of course to satisfy all these conditions many times is not possible, and to make use of some hierarchical rules is necessary.

RESULTS AND DISCUSSION

About 60,000 cases were calculated using the built neural network model in order to map a grid of process conditions from which the most convenient ones were selected. The range of each input variable was divided into an equal number of intervals.

Tables 3 - 6 show two examples of cases for this optimization procedure. Tables 3 and 5 show the selected conditions for feed specifications according to the domain of these variables considered in the neural network modeling as well the restrictions established for independent variables.

From a qualitative point of view, some conditions lead to a less expensive operation; for example, a low reflux flow is better than a high one, as well as low reboiler steam with respect to saving energy. A low temperature is better than a higher one, not due to energy consumption but mainly because higher temperature may favor dimerization reactions, which increase IP loss and this situation is not desirable.

From Tables 4 and 6, numerous cases that satisfy the established restriction are observed, but the final solution of each case studied presents only a few operational conditions that satisfy the criteria of a less expensive operation.

Table 3. Optimization specifications – Case 1

Feed specification		Operational Restriction	
x1(1)	14.00	IP losses	< 5.5%
x1(3)	17.99	CPD in reactor [x2(4)]	< 3.5%
x1(4)	5.95		
x2(1)	= x1(1)	IP in distillate [y2(2)]	> 23.0%
x2(3)	= y1(2)	CPD in distillate [y2(3)]	< 3.0%
x2(4)	=x1(4)-y1(3)		
x3(1)	= y2(5)		
Cases which satisfy the restrictions		4212	
Energy minimization: lowest values for		x1(2)	x3(2)
		x2(5)	x3(3)
		x2(7)	x3(4)

Table 4. Qualitatively more economic situations – Case 1

Final Solutions																		
X1(2)	Y1(1)	Y1(2)	X2(2)	X2(4)	X2(5)	X2(6)	X2(7)	Y2(1)	Y2(2)	Y2(3)	Y2(4)	Y2(5)	X3(2)	X3(3)	X3(4)	Y3(1)	Y3(2)	Y3(3)
75.00	78.76	17.01	63.97	3.41	46.17	59.77	9.57	7.42	35.40	2.63	89.76	6.30	56.60	3.07	1.13	2.07	4.05	50.76
75.00	78.76	17.01	63.97	3.41	46.17	59.77	9.57	7.42	35.40	2.63	89.76	6.30	56.60	3.07	1.13	2.51	3.89	47.96
75.00	78.76	17.01	63.97	3.41	46.17	59.77	9.57	7.42	35.40	2.63	89.76	6.30	56.60	3.07	1.13	2.96	3.58	47.34
75.00	78.76	17.01	63.97	3.41	46.17	61.20	9.57	6.97	37.27	2.12	86.79	6.30	56.60	3.07	1.13	2.07	4.05	50.76
75.00	78.76	17.01	63.97	3.41	46.17	61.20	9.57	6.97	37.27	2.12	86.79	6.30	56.60	3.07	1.13	2.51	3.89	47.96
75.00	78.76	17.01	63.97	3.41	46.17	61.20	9.57	6.97	37.27	2.12	86.79	6.30	56.60	3.07	1.13	2.96	3.58	47.34

Table 5. Optimization specifications – Case 2

Feed specification		Operational Restriction	
x1(1)	13.50	IP losses	< 6.5%
x1(3)	21.54	CPD in reactor [x2(4)]	< 3.5%
x1(4)	6.79		
x2(1)	= x1(1)	IP in distillate [y2(2)]	> 30.0%
x2(3)	= y1(2)	CPD in distillate [y2(3)]	< 2.5%
x2(4)	= x1(4)-y1(3)		
x3(1)	= y2(5)		
Cases which satisfy the restrictions		5265	
Energy minimization: lowest values for		x1(2)	x3(2)
		x2(5)	x3(3)
		x2(7)	x3(4)

Table 6. Qualitatively more economic situations – Case 2

Final Solution																		
x1(2)	y1(1)	y1(2)	x2(4)	x2(2)	x2(5)	x2(6)	x2(7)	y2(5)	y2(1)	y2(2)	y2(3)	y2(4)	x3(2)	x3(3)	x3(4)	y3(1)	y3(2)	y3(3)
74.00	81.84	20.21	3.45	63.97	46.17	59.77	9.57	6.30	6.72	37.33	2.27	87.19	56.60	3.07	1.13	2.07	4.05	50.76
74.00	81.84	20.21	3.45	63.97	46.17	59.77	9.57	6.30	6.72	37.33	2.27	87.19	56.60	3.07	1.13	2.51	3.89	47.96
74.00	81.84	20.21	3.45	63.97	46.17	59.77	9.57	6.30	6.72	37.33	2.27	87.19	56.60	3.07	1.13	2.96	3.58	47.34
74.00	81.84	20.21	3.45	63.97	46.17	61.20	9.57	6.30	6.75	37.83	2.02	86.71	56.60	3.07	1.13	2.07	4.05	50.76
74.00	81.84	20.21	3.45	63.97	46.17	61.20	9.57	6.30	6.75	37.83	2.02	86.71	56.60	3.07	1.13	2.51	3.89	47.96
74.00	81.84	20.21	3.45	63.97	46.17	61.20	9.57	6.30	6.75	37.83	2.02	86.71	56.60	3.07	1.13	2.96	3.58	47.34
78.20	87.79	20.16	2.92	63.97	46.17	59.77	9.57	6.30	6.75	37.17	2.26	87.37	56.60	3.07	1.13	2.07	4.05	50.76
78.20	87.79	20.16	2.92	63.97	46.17	59.77	9.57	6.30	6.75	37.17	2.26	87.37	56.60	3.07	1.13	2.51	3.89	47.96
78.20	87.79	20.16	2.92	63.97	46.17	59.77	9.57	6.30	6.75	37.17	2.26	87.37	56.60	3.07	1.13	2.96	3.58	47.34
78.20	87.79	20.16	2.92	63.97	46.17	61.20	9.57	6.30	6.82	37.80	2.02	86.71	56.60	3.07	1.13	2.07	4.05	50.76
78.20	87.79	20.16	2.92	63.97	46.17	61.20	9.57	6.30	6.82	37.80	2.02	86.71	56.60	3.07	1.13	2.51	3.89	47.96
78.20	87.79	20.16	2.92	63.97	46.17	61.20	9.57	6.30	6.82	37.80	2.02	86.71	56.60	3.07	1.13	2.96	3.58	47.34

CONCLUSION

The optimization procedure based on a grid search of the problem using the neural network as an industrial plant model has an advantage of mapping all the solutions.

This procedure allows one to define the domain where the optimal solution is located. Quantitative and qualitative solutions can be promptly analyzed. To have an explicit objective function in order to get optimal operation points is not necessary since it is possible to work only with inequalities.

Acknowledgements

The authors gratefully acknowledge FAPESP for their financial support and COPENE Petroquímica do Nordeste for providing the industrial data used in this work.

References:

- 1 Nascimento, C.A.O., Giudici, R., Guardani, R. (2000), 'Neural Network Based Approach for Optimization of Industrial Chemical Processes', *Comput. Chem. Eng.*, 24, 2303.
- 2 Seider, W.D., Seader, J.D. & Lewin, D.R., (1999) *Process Design Principles: Synthesis, Analysis, and Evaluation*, John Wiley & Sons, Inc.
- 3 Nascimento, C.A.O., Giudici, R. (1998), 'Neural Network Based Approach for Optimization Applied to an Industrial Nylon-6,6 Polymerization Process', *Comput. Chem. Eng.*, 22, S595.
- 4 Willis, M.J., Di Massimo, C., Montague, G.A., Tham, M.T., Morris, A.J. (1991), 'Artificial Neural Networks in Process Engineering', *IEE Proceedings-D*, 138 (3), 256-266.
- 5 Schenker, B. and Agarwal, M. (1996), 'Cross-Validated Structures Selection for Neural Networks', *Comp. Chem. Engng.*, 20 (2), 175-186.
- 6 Alves, R. M. B. & Nascimento, C. A. O. (2001), 'Gross Error Detection of Industrial Data by Neural Network and Clusters Techniques', submitted to *ENPROMER 2001*, Santa Fé, Argentina.
- 7 Pollard, J. F., Broussard, M. R., Garrison, D. B. & San, K. Y. (1992), 'Process Identification Using Neural Networks', *Comp. Chem. Eng.*, 16(4), 253-270.

MODELING OF AN INDUSTRIAL PROCESS FOR ISOPRENE PRODUCTION USING A NEURAL NETWORK BASED APPROACH

R. M. B. Alves and C. A. O. Nascimento

LSCP-Process Control and Simulation Laboratory
Department of Chemical Engineering, Polytechnic School, University of São Paulo
Av. Prof. Luciano Gualberto, n. 380, trav. 3, 05508-900, São Paulo - SP, Brazil
Telephone number: + 55 11 3091-2237; Fax number: + 55 11 3813-2380
e-mail: rita@lscp.pqi.ep.usp.br; oller@usp.br

Abstract. This article describes the application of a three-layer feed-forward neural network (NN) to model an industrial plant for Isoprene production. The purpose of this plant is to produce high purity Isoprene for obtaining synthetic and thermoplastic rubber from a C₅ cut arising from a pyrolysis gasoline unit. The basic idea is the modeling of this entire chemical process using NNs, which replace phenomenological models. Since most chemical processes are non-linear, complex and multivariable systems, NN appears as a technique of interest due to its capability of learning the system without knowledge of the physical and chemical laws that govern it. However, success in obtaining a reliable and robust NN depends strongly on choosing the process variables involved, as well as the quality of the available data set and the domain used for training purposes. Since NNs are able to extract information from plant data in an efficient manner, for this work the neural network model was built directly from historical plant data, which were collected every of 15 min during a period of one year. The modeling using NN can be carried out in a global way or by parts. This last approach was the chosen one because, from this way, it is possible to obtain information about intermediate streams. This step consisted of establishing optimal neural network structures to represent each process unit individually. Then, the global model was built, by interconnecting each individual model through information in common among them. Comparisons between the model's prediction and the experimental data were performed and reasonable results were achieved from an industrial point of view. This new approach provides comprehensive information for an engineer's analysis than the conventional procedure and the use of NN methodology is promising for several industrial applications, such as data analysis, process modeling, simulation and optimization.

Keywords: Neural Network, Process Modeling, Isoprene Production Unit

INTRODUCTION

Process analysis, or simulation, is the tool chemical engineers use to interpret flowsheets, to locate malfunction, and to predict the performance of processes. The heart of analysis is the mathematical model, a collection of equations that relate the process variables, such as stream temperature, pressure, flow rate, and composition, to surface area, valve settings, geometrical configuration, and so on.

The steady-state simulators solve the unknown variables, given the values of certain quantities, i.e., given the detailed process flowsheet for the base-case design, or a process flowsheet after the task integration step in the process synthesis, or even an incomplete flowsheet after one of the earlier steps, it is often possible to use a process simulator to solve many unknown temperatures, pressures, and flow rates in the steady-state. For an existing process, analysis using a process simulator is performed routinely to study potential changes in the operating conditions or the possibility of a retrofit to improve its profitability.

There are several levels of analysis. In order of increasing complexity, they involve: material balances, material and energy balances, equipment sizing, and profitability analysis. Additional equations are added at each level. New variables are introduced, and the equation-solved algorithms become more complicated. Fortunately, most chemical processes involve conventional process equipment: heat exchangers, pumps, distillation columns, absorbers, and so on. For these process units, the equations do not differ among chemical processes. The physical and thermodynamic properties and chemical kinetics constants differ, but not the equations. Hence, it is possible to prepare one or more equation-solving algorithms for each process unit to solve the material and energy balances and to compute equipment sizes and costs. A library of subroutines or models that automates such equation-solving algorithms is at the heart of process simulators. The steady-state models in simulators simulate only the steady-state operation of process units (operation subroutines) and estimate the sizes and costs of process units (cost subroutines). Two other types of subroutines are used to perform optimization and/or converge recycle computations (convergence subroutines) and to alter the equipment parameters (control subroutines).

When using the process simulators, it is important to recognize that, with some expectations, most streams are comprised of chemical species that distribute within one or more solution phases that are assumed to be in phase equilibrium. Then it is important, when using process simulators, to understand how they apply the theory of phase equilibrium in modeling streams as well as in some vapor-liquid equipment.

As said above, commercial simulators are important tools largely used on real plants and new design, but an important prerequisite for process modeling is the availability of a valid and representative mathematical model that will be able to describe the system accurately. When available, a first-principle based model could be a natural choice. However, in most chemical engineering problems, such models are non-linear, complex, multivariate systems and computationally time-consuming. Consequently to use commercial simulators is a very hard and difficult task, mainly if fundamental knowledge about the process under consideration is missing or there is a lack of thermodynamics data. In this sense some works have been proposed to use artificial neural network as a substitute for first-principle models (e.g. Nascimento and Giudici, 1998; Guardani et al., 2001; Sternowsky et al., 2002).

Neural networks (NNs) have been claimed to be a universal non-linear approximator and their application in the field of chemical engineering has grown rapidly. NNs replace the phenomenological models due their capability of learning the system without knowledge of the physical and chemical laws that govern it, and are not computationally time-consuming (Nascimento and Giudici, 1998). The advantage of the NN model over other empirical approaches will depend strongly on the degree of non-linearity of the process. However, the success in obtaining a reliable and robust NN depends strongly on choosing the process variables involved, as well as the quality of the available data set and the domain used for training purposes. It is possible to obtain a representative NN model based on historical input/output data and/or designed plant experiments.

NEURAL NETWORK MODELS

In the context of chemical engineering applications, neural networks are of particular interest as predictive models (Nascimento et al., 1994, 1999 and 2000) and for pattern

recognition (Nascimento et al., 1997). Moreover, neural networks have become a widely used tool in areas such as process analysis, simulation, and control, primarily due to their inherent advantages such as adaptability, non-linearity, and fault tolerance especially in combination with the relatively simple way of use. Neural networks possess the ability to 'learn' the behavior of the process without actually specifying the physical and chemical laws that govern the system. In addition, neural network models can simultaneously employ continuous and discrete input variables (Migliavacca et al., 1999). The success in obtaining a reliable and robust network strongly depends on the choice of appropriate input or process variables, as well as the available set of data and the domain used for training purposes.

The neural network employed in this study is a three-layer feed-forward network. This type of network, in which information propagates in only one direction, is particularly useful for steady-state modeling. The architecture of a typical NN is shown in Figure 1. The nodes in different layers of the network represent "neuron-like" processing elements. The number of neurons in input and output layers depends on the respective number of inputs and outputs being considered. On the other hand, hidden layers may vary, depending on the user specification, and its structure defines the feed forward NN topology. The neurons in the input layer do not perform data processing functions; they merely provide a means by which scaled data is introduced into the network. These signals are then "fed forward" through the net via the connections, through hidden layers, and eventually to the final output layer. Each interconnection has associated with it a weight that modifies the strength of the signal flowing along that path. Thus, the input to each neuron is a weighted sum of the outputs from neurons in the previous layer. The output of each node is obtained by passing the weighted sum through a non-linear operator or activation function. Although other forms of activation functions can be adopted, in this work, a sigmoid function was used, because it can be considered as an universal approximator and is able to represent non-linear relationships. Moreover, this function provides the network with the ability to fit convex and concave multivariable functions (Cybenko, 1989). Selection of an optimal network structure is crucial for network performance. For example, a network with too few hidden nodes cannot represent the system satisfactorily and a network with too many hidden nodes overlearns particular runs presented in the training set.

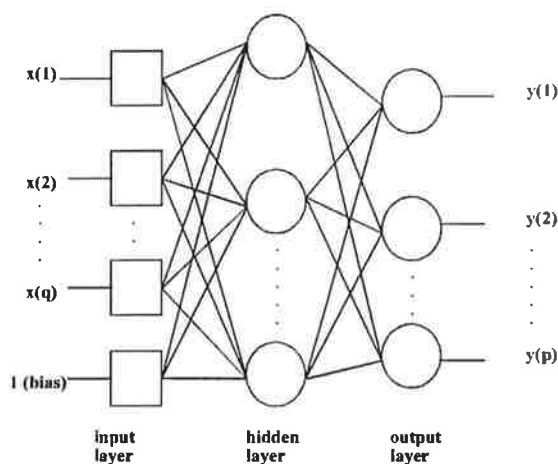


Fig. 1. Structure of a feed forward neural network

Once the network topology is specified, a set of input-output data is used to train the network; i.e., to determine appropriate values for the weights associated with each interconnection. The data are propagated forward through the network to produce an output that is compared with the corresponding output in the data set, hence generating an error. This error is minimized by making changes to the weights, and may involve many passes through training data set. For a given topology, the magnitude of the weights defines the network characteristics and the structural properties of the model. Thus, an NN has the capability to represent complex systems whose structural properties are unknown (Willis et al., 1991).

The typical network design consists of training various network structures with one part of the data, the so-called training or learning set, which is used to train the network, i.e., to fit the data and should preferably contain the extremes of all the input and output variables; an their performance is evaluated afterwards with the remaining part of the data, the so-called test set, which is used to verify the prediction capability of the trained network, i.e., to validate the model. Both sets can be made large enough to ensure satisfactory learning of the network weights and accurate evaluation of the network performance (Schenker and Agarwal, 1996). In training a network, the objective is to find an optimum set of weights. When the number of weights is larger than the number of available data points, the error in fitting the non-trained data initially decreases, but then increases as the network becomes over trained. In contrast, when the number of weights is much smaller than the number of data points, the overfitting problem is not crucial.

The most extensive algorithm for the learning phase is the back-propagation algorithm, which is a generalization of the steepest descent method (Haykin, 1994; Rummelhart and J. McClelland, 1986). Since the NN model fitting is essentially a mathematical adaptive regression in which phenomenological considerations are not used, much care must be taken to validate its representation of the physical process and to prevent overfitting. A simple measure of the quality of the fitting for a given NN is based on comparisons between calculated values and experimental data from the test set (not used in fitting the NN).

The computer programs for data preparation, NN fitting, and modelling used in this work were developed at the LSCP (Process Simulation and Control Laboratory) of the Chemical Engineering Department of the University of São Paulo

PROCESS DESCRIPTION

The system studied is the Isoprene Production Unit from COPENE, the largest Brazilian petrochemical plant.

The objective of the Isoprene Extraction Unit is to yield high purity isoprene from a C₅ cut of pyrolysis gasoline. The isoprene produced is to be used to obtain synthetic and thermoplastic rubber. The isoprene production is performed by an extractive distillation process by using a mixture of Acetonitrile (ACN), Water, Allyl Alcohol and Acetone at a specified proportion as solvent. The process of isoprene production can be divided into five sections: feed preparation, extractive distillation, solvent recovery, sulfur removal and fractionating. The system consists basically of a dimerization reactor and a separation column train. At the feed preparation section, the unit feed stream flows through a tubular

reactor to thermally dimerize cyclopentadiene (CPD) into dicyclopentadiene (DCPD). The reactor effluent is fed to a distillate column for removal of heavies. This first column – Heavies Removal First Column - separates one stream rich in isoprene, as the distillate, from a heavy cut, rich in DPCD and piperylenes, as the bottom product. This bottom stream feeds another column – Depentanizer Column - to separate the heavy cut into two streams: one rich in DCPD and the other rich in piperylenes, both of them used as raw material in resin manufacture. The distillate product of the first column is the feed stream of the isoprene extraction section, which uses a mixed solvent to separate high purity isoprene from the other compounds of the C₅ stream. The first extractive distillation column is physically divided into two sections. The solvent is fed into stage 15 of the first section, and the distillate from the column for removal of heavies is fed into stage 13 of the second section. Paraffins are obtained as a top product and diolefins, whose main component is the isoprene, are obtained together with the solvent as a bottom product. The bottom product of the first column of extractive distillation is fed to solvent stripper column, into stage 16. The top stream of this column is fed into the last stage of the second extractive distillation column. In this column the solvent is fed into stage 11. The ratio solvent/feed for this column is less than those used in the first extractive distillation column in order to get, as the product top, a stream rich in isoprene, that is further washed with water in a liquid-liquid extraction column – Product Washing Column - in order to remove the entrained solvent completely.

The bottom stream from the second extractive distillation column feeds also the solvent stripper column, where the bottom recovers all the solvent used in the both extractive distillation columns. After heat recovery, this recycle stream is fed into the extractive distillation columns of the process.

The side stream of the solvent stripper column is sent to one packed column – the CPD Removal Column - where the top removes part of the butene-2, CPD, piperylenes and 1,4-pentadiene retained with the solvent. The bottom of this column is solvent quite pure, which is returned back to the solvent stripper column.

All the raffinate streams of the extraction section are washed with water in a liquid-liquid extractive column - C₅ Washing Column - in order to recover the entrained solvent. The bottom streams of the two washing columns, that contain the extracted solvent, are fed into a distillation column – Solvent Recovery First Column – in order to separate water, which has practically no ACN, as the bottom stream from the stream rich in ACN as the top stream. This stream contains more water than the rate specified for the circulating solvent and returns back to the solvent circuit. The water control in this stream allows the system to keep the water concentration at the desired value.

The hydrocarbons from the top of C₅ washing column are sent to the C₅ raffinate vessel and then they go on to the raffinate spheres, from where they are sent to mono-hydrogenation at the other unit and co-cracking in the pyrolysis furnace.

Isoprene Crude from the product-washing column is fed to the sulphur removal system. After that, the final isoprene purification is performed in two fractionators: the first one is for removal of lights and the second one is for removal heavies, and at the top of the last one, isoprene product with purity higher than 99,9% is obtained.

Figure 2 shows the Isoprene Unit schematically.

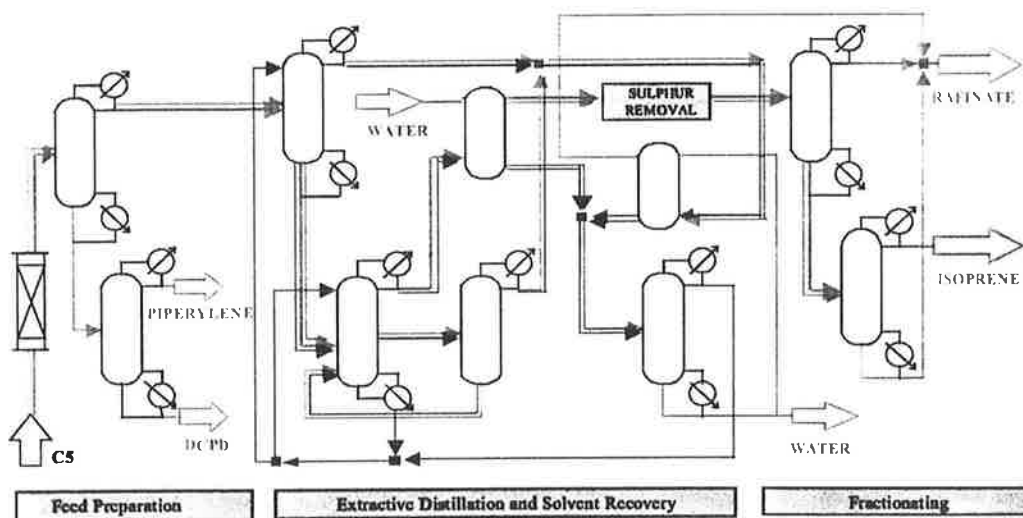


Fig. 2 Isoprene Unit

PROCESS MODELING

Methodology

As said previously, modeling through neural networks consists, in general, in determining weights associated with each data input. These weights are values that minimize the quadratic error between the value calculated by neural networks and the experimental data provided as output. Thus, it is very important to keep in mind the procedures to be performed in order to adjust models by neural networks. The steps involved are described below:

1. Analysis of the problem to be modeled by neural networks

This step involves the consideration of each variable, which characterizes the process, including the binaries ones, since the training procedure is similar to the learning process of the human mind, lack of important information leads the network to determine incomplete non-linear relationships in the same way and its predictive capability will be lost. The use of redundant information should also be avoided.

At this step, operational, environmental, safety restriction, product specifications, maximum and minimum limits and sensitivity of all the variables involved should be considered.

2. Data Collection

Quality and quantity of the process data are relevant. Debatable confidence of the values could prejudice the network training and consequently the prediction step will show low quality.

The analysis of the process was undertaken by using a one-year database. The primary database consisting of about 34500 sets of observations of 244 variables

3. Statistical Analysis and Data Reconciliation

The neural network modeling was done directly from industrial data. However presence of noise and gross errors, commonly found in industrial processes would disturb the modeling and lower its quality. Therefore, for this work, the first step was to prepare the industrial data to a model fitted to detection and elimination of data containing gross errors ("outliers") and systematic errors.

This step is extremely important since the quality of the data has direct influence on the quality of the fitting performed by the neural networks. Knowledge of the process, statistical procedures and others based in the first principles equations are used.

The available monitoring variables from the industrial process studied (Isoprene Production Unit) were collected every 15 minutes. According to the average time considered for the data treatment, data fluctuation could be incorporated in the results. Many times, this could lead to unreliable information. In cases of errors with the measurement instruments over a long period of time, the average reflects this error.

The data treatment can be divided into the following sections:

✓ Gross Error Detection

At this step the following data were evaluated and eliminated: null and negative values, values with different magnitudes, possible flat lines like those at the instrument's limits (wind-up measurement) and abrupt changes of the variables along the time line. This analysis was carried out through observations of the variables as a time function, by verification of their minimum, maximum and mean values. The knowledge of the process also allowed the elimination of some points based on possible process values or the acceptable operational range for the corresponding variable.

Another tool used to detect outliers was cluster analysis. Cluster analysis is an analytical technique for developing meaningful subgroups of individuals or objects. It is based on the similarity principle among several data sets. For this work, a data set was formed by the input and output variables chosen for each process unit, corresponding to information from one operation register. It is expected that for a series of similar input variables, the process must yield similar output variables (dependent variables). When a different input or output variable is observed among a series of similar data, the corresponding data set may be rejected.

✓ Systematic Errors Detection

At this step, first principle procedures were used in order to detect systematic errors. Knowledge of the process is also important at this step in order to evaluate these kinds of error. Once carried out, global and

components material balances were possible to identify some distortion in the final results. At this point it was very informative to make a graphical representation of a frequency distribution. Based on this procedure, it was possible to correct systematic errors by shifting the mean. Another criterion used at this step was the comparison between the plant data and the laboratory data analysis, which was more reliable. This procedure allowed identification of possible errors in the measurement instruments at certain periods of plant operation and correction of the wrong values.

✓ Establishment of Steady-State Operation

The higher frequency of data collected allowed identifying periods of steady state operation. The criterion adopted was a constant feed flow fluctuation of 0.2-0.3 t/h during a period of 48 to 72 hours.

- ✓ After this initial analysis, the resulting data set was used to train the neural network. At this step, all data were included in the training data set and during the training process few iterations were performed. It was observed that some points and groups of points were not well adjusted. The identification of these groups (or points) is an indication of consistency problems, gross and systematic errors not identified in the previous step of the procedure. To decide if these points must or must not be eliminated, they were analyzed by according their significance to the process and after this some statistical analyses as clusters analysis and hypothesis tests for means were used. This procedure was repeated until anomalous points did not exist any more.

More details of these procedures can be seen in Alves and Nascimento (2001) and Alves et al. (2002).

4. Modeling of the System

With the system input and output previously selected at the treatment of the historical data step at the section 3, neural networks can model the process in distinct ways, among them: networks exist that model the whole process at once; networks that model the process by parts and then proceed with a connection among the several networks. In this case, the option was to divide the process in several subsystems, each one representing a given unit operation (reactor, distillation column, and so on). This methodology has as objective:

- ✓ To simplify the modeling by decreasing the neural network order.
- ✓ To make the plant data consistency better in each subsystem. Measurement instruments errors in one variable in one subsystem do not prejudice, most of time, the information to other subsystems.
- ✓ An important part of the process recycle stream is explicit, which facilitate the technical and economical analysis of the process.

In this way, for each subsystem, every process input and output measured variable could be analyzed concerning its importance and consistency.

After the modeling of the individual neural network, which represents each subsystem, the main focus is to interconnect sequentially these networks in order to generate data with the objective of building a response surface for all involved variables, as well as analyzing the behavior of the most important process variables. Thus a “black box” was built, as shown as an example in Figure 3, containing all the networks interconnected by recycle and sequential streams.

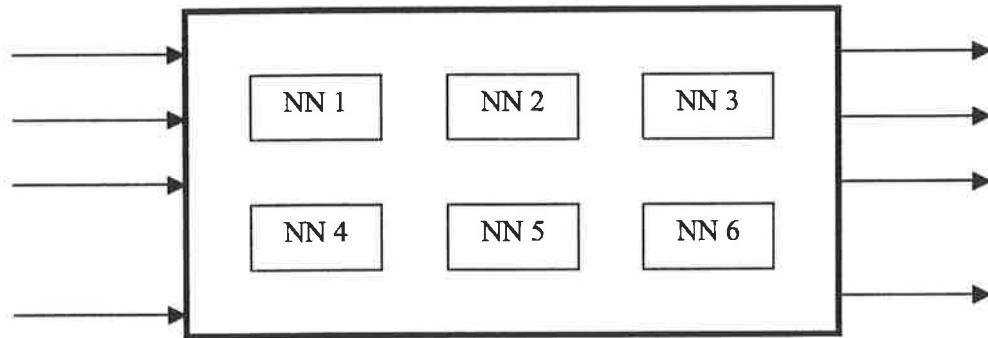


Fig. 3. “Black box” Scheme

The convergence procedure of the integrated system was performed interconnecting all the streams subsequently at the corresponding unit operation or subsystem, in order to reproduce the industrial plant. The algorithm used for mathematical convergence was direct substitution with dumping factor.

5. Neural Network Training

After the plant data analysis, the process modeling was performed as explained previously. This step consisted of establishing an optimal neural network structure to represent each process unit individually, i.e., the number of neurons in the hidden layer, the number of iterations for the learning process, the residual value and the learning and test sets were determined.

Initially, the industrial data was divided into two sets: a learning set and a test set. The test set contained variables within the range of the learning set variables. Thus, no extrapolation takes place in the model in comparison with the test set data.

Then, the NN was run with several different specifications, i.e., varying the number of neurons in the hidden layer and the number of data presentations (iterations), in order to select the best one. In this way several networks with different errors are generated. To avoid an overfitting of the model, choosing a minimum neuron number in the hidden layer (NH), as well as an adequate number of training cycles that accomplishes the lowest error in the test set is recommended (Pollard, 1992). Figure 4 shows an example of this procedure, i.e., the minimum error for the test set as a function of the

number of neurons and presentations. A minimum value for error is reached for NH equals 9 and about 30000 presentations can be observed.

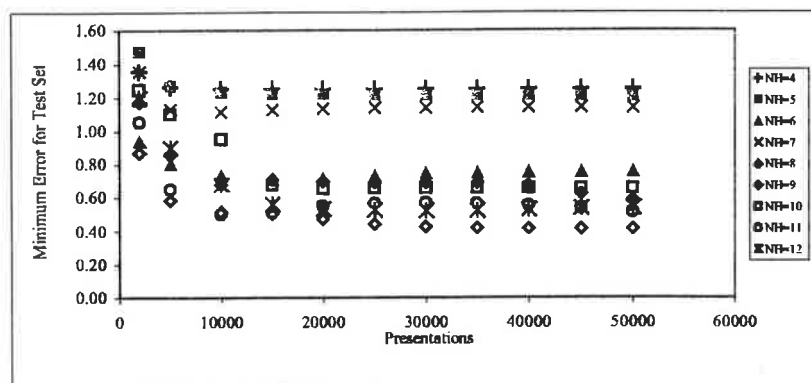


Fig. 4. Minimum error for the test set

The phase of learning or training is, basically, a problem of parameters estimation and, in general, the same difficulty associated to conventional optimization, such as, convergence, local minima and time consuming can be found.

6. Validation of the training neural network

Once the best neural network that represents the output variables was selected, tests of consistency were carried out in order to validate the network. The first evaluation is the analysis of dispersion by comparison between the calculated values and the measured data. Then, the normal distribution of errors between the calculated and measured values is verified. This information evaluates the quality of the fit and the normal distribution demonstrates the equilibrium of frequency of the values located around the medium zero.

RESULTS and DISCUSSION

The Isoprene Process Unit was divided into 10 subsystems in order to carry out the modeling as described previously in the item Adjustment and Modeling of the System:

- ✓ Subsystem 1 – Reactor
- ✓ Subsystem 2 – Heavies Removal First Column
- ✓ Subsystem 3 – Depentanizer Column
- ✓ Subsystem 4 – First Column of Extractive Distillation
- ✓ Subsystem 5 – Second Column of Extractive Distillation and Stripper Column Solvent
- ✓ Subsystem 6 – C₅ Washing Column and Product Washing Column

- ✓ Subsystem 7 – CPD Removal Column
- ✓ Subsystem 8 – Solvent Recovery First Column
- ✓ Subsystem 9 – Lights Removal Column
- ✓ Subsystem 10 – Heavies Removal Second Column

As an example, Table 1 shows the variables selected and their applicability range for the subsystems 1, 2 and 3 that constitute the feed preparation section:

Table 1. Process variables and their application range

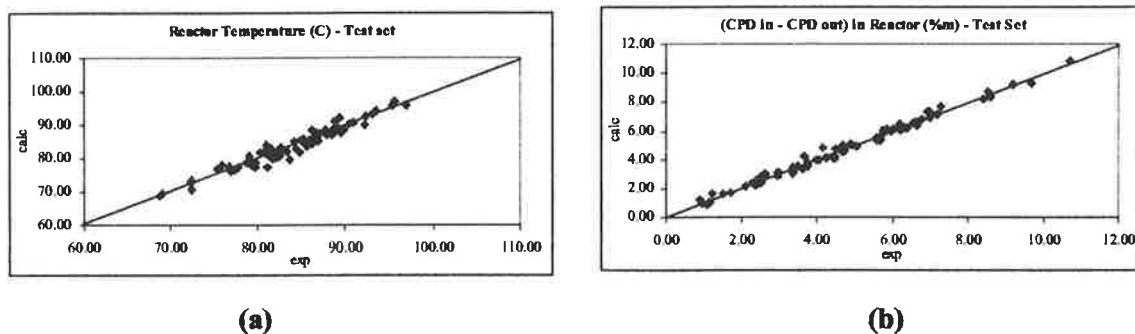
Process Unit	Variables	Notation	Unit	Maximum	Minimum
Subsystem 1	Feed flow	x1(1)	t/h	14.51	9.49
	Feed temperature	x1(2)	°C	79.0	62.0
	IP in feed	x1(3)	%weight	23.68	13.48
	CPD in feed	x1(4)	% weight		
	Reactor temperature	y1(1)	°C	100,5	68,3
	IP in reactor	v1(2)	%weight	21,64	12,27
	CPD feed–CPD reactor	v1(3)	%weight	13,59	21,64
Subsystem 2	Feed flow	x2(1)	t/h	14.51	9.50
	Feed temperature	x2(2)	°C	66.5	62.6
	IP in feed	x2(3)	%weight	21,48	12,27
	CPD in feed	x2(4)	%weight	4,74	2,01
	Reflux flow	x2(5)	t/h	66,3	35,1
	Reflux temperature	x2(6)	°C	61,20	56,9
	Reboiler Steam	x2(7)	t/h	11,90	8,41
	Distillate flow	y2(1)	t/h	9,10	5,92
	IP in distillate	y2(2)	%weight	38,72	23,55
	CPD in distillate	y2(3)	%weight	4,43	1,87
	Bottom temperature	y2(4)	°C	103,8	88,6
	Bottom flow	y2(5)	t/h	6,30	2,75
	Subsystem 3	Feed flow	x3(1)	t/h	6.58
Feed temperature		x3(2)	°C	61,9	52,9
Reflux flow		x3(3)	t/h	4,50	3,99
Reboiler Steam		x3(4)	t/h	1,45	1,12
Distillate flow		y3(1)	t/h	2,49	1,55
Bottom flow		y3(2)	t/h	4,28	1,58
Distillate temperature		y3(3)	°C	52,6	48,4

For the present work, the selected neural network configurations are shown in Table 2 for the three subsystems shown in the Table 1:

Table 2. Neural network configurations

	Neurons in hidden layer	Presentations	Number of data set in learning set	Number of data set in test set
Subsystem 1	9	30000	165	78
Subsystem 2	12	30000	156	63
Subsystem 3	13	45000	164	75

Figures 5a and 5b show some examples of the comparison between measured and calculated output variable values for the test data set. Good agreement between modeled and measured values can be observed and the uniform distribution indicates absence of any tendency on fitting.



Figs. 5a,b. Comparison between experimental and calculated values

The global model is complex due to the great number of variables and recycles involved. The global model generated and the structure of convergence employed (direct substitution with dumping factor) show good efficiency in representing the process.

Once the neural network was trained and the global model was well adjusted, the model could be used to simulate the process with great speed. Millions of simulations can be performed in seconds, which allows its use in procedures of process optimization, process control, on-line process control, analysis of process restriction, sensitivity analysis of variables, virtual analyzers, among other applications.

CONCLUSIONS

The global model built as a sequence of NN, interconnected by recycles and intermediate variables, shows good results, which allows a deep understanding of the behavior of the process variables and consequently generates an important tool either for a process supervisor and also for training people who work in plant operation. In this way it was possible to simulate this industrial process since because of the lack of thermodynamic data it was impossible to use conventional process simulators.

Among several advantages that modeling by neural networks bring, one of the most important is the detection of measurement error, which probably indicates what instrument presented some problem in a given period.

Acknowledgements

The authors gratefully acknowledge FAPESP for its financial support and COPENE Petroquímica do Nordeste for providing the industrial data used in this work.

References

- Alves, R.M.B. and Nascimento, C.A.O. (2001), "Gross Error Detection of Industrial Data by Neural Network and Clusters Techniques", Proceedings of ENPROMER 2001, Santa Fé, Argentina.
- Alves R. M. B, Nascimento C. A. O. and Carneiro, W.L. (2002), "Analysis and Detection of Outliers and Systematic Errors from an Industrial Data Plant", submitted to *FOCAPO 2003*, Radisson, Coral Springs, Florida, USA.
- Cybenko, G. (1989), *Mathematics of Control, Signals and Systems*, 2nd ed., 303.
- Guardani, R., Onimaru, R.S. and Creso, F.C.A. (2001), "Neural Network Model for the On-Line Monitoring of a Crystallization Process", *Brazilian Journal of Chemical Engineering*, v. 18(3), 267-275.
- Haykin, S. (1994). *Neural Networks – A Comprehensive Foundation*. New York: Macmillan College Publishing Company.
- Migliavacca, S.C.P., Nascimento, C.A.O. and Rodrigues, C. (1999), "Use of Neural Network for the Simulation of a Gas Centrifuge", *J. Nuclear Science. & Technology*, 36(4), 364-370.
- Nascimento, C.A.O., Oliveros, E. and Braun, A.M., "Neural Network Modeling for Photochemical Process", *Chem. Engineering and Processing*, 33, 319-324.
- Nascimento, C.A.O., Guardani, R. and Giulietti, M. (1997), Use of Neural Networks in the Analysis of Particle Size Distribution by Laser Diffraction, *Powder Technol.*, v.80 (1), 89-94.
- Nascimento, C. A. O. and Giudici, R., (1998), "Neural Network Based Approach for Optimization Applied to an Industrial Nylon-6,6 Polymerization Process", *Comput. Chem. Eng.*, 22, S595-600.
- Nascimento, C.A.O., Giudici, G. and Scherbakoff, N. (1999). "Modeling of Industrial Nylon-6,6 Polycondensation Process. Part 2: Neural Networks and Hybrid Models", *J. Applied Polymer Science*, 72, 905-912.
- Nascimento, C.A.O., Giudici, R., Guardani, R. (2000), "Neural Network Based Approach for Optimization of Industrial Chemical Processes", *Comput. Chem. Eng.*, 24, 2303-2314.
- Pollard, J. F., Broussard, M. R., Garrison, D. B. and San, K. Y. (1992), "Process Identification Using Neural Networks", *Comp. Chem. Eng.*, 16(4), 253-270.
- Rummelhart, D. E. and McClelland, J. L. (1986). *Parallel Distributed Processing* (Vol. 1), Cambridge, MA: MIT Press
- Schenker, B. and Agarwal, M. (1996), "Cross-Validated Structures Selection for Neural Networks". *Comp. Chem. Engng.*, 20 (2), 175-186.
- Sternowsky, S, Schulte G., Guardani R. and Nascimento C.A.O. (2002), "Experimental Study and Neural Network Modeling of the Ligament Disintegration in Rotary Atomization", *Atomization and Sprays*, v. 12, 1-15.
- Willis, M.J., Di Massimo, C., Montague, G.A., Tham, M.T., Morris, A.J. (1991), "Artificial Neural Networks in Process Engineering", *IEE Proceedings-D*, 138 (3), 256-266.

NEW APPROACH FOR THE PREDICTION OF AZEOTROPY IN BINARY SYSTEMS

R.M.B. Alves^{*1}, F.H. Quina² and C.A.O.Nascimento¹

¹LSCP- Process Control and Simulation Laboratory
Department of Chemical Engineering, Polytechnic School, University of São Paulo
Av. Prof. Luciano Gualberto, n. 380, trav. 3, 05508-900, São Paulo – SP, Brazil
²Institute of Chemistry, Univ. of São Paulo, CP 26077, São Paulo 05513-970, Brazil

Abstract. The ability to predict whether a given mixture will form an azeotrope and which is the azeotropic point is essential in any separation process. Since the determination of azeotropy strictly from experiment can be expensive and time-consuming, the availability of computational methods for predicting azeotropes would not only reduce this cost, but also save time by narrowing the experimental search space. The objective of the present work is to predict azeotropy based only on properties of the individual components of the mixture, i.e., without any explicit consideration of non-ideality of mixing of the components in the liquid phase. Neural networks are used to correlate available experimental information on azeotropy in order to predict whether or not a given binary mixture should be expected to form an azeotrope and, if it exists, which is its temperature and composition at atmospheric pressure. A three-layered feed-forward neural network was used and trained with the back-propagation algorithm. The model results provide rather good predictions of the occurrence (or absence) of azeotropy and the correspondent azeotropic conditions in a wide variety of binary systems. The model fails to make a clear prediction regarding azeotropy in only a relatively small number of situations corresponding to structurally homologous molecules that exhibit quite distinct azeotropic behavior.

Keywords: Neural Network; Azeotropy; Azeotropic prediction

1. Introduction

The ability to predict whether a given mixture will form an azeotrope, including the conditions required for such azeotrope formation and the resulting azeotropic compositions, is essential in any separation process. Moreover, an understanding of azeotropy is important for the successful selection of entrainers in azeotropic distillation, of auxiliary materials in extractive distillation and of the appropriate solvents in extractions.

Because of the crucial importance of azeotropy in the design and synthesis of distillation processes, a large quantity of experimental data on azeotropic behavior has been compiled over the years (Horsley, 1973, Gmehling et al., 1994). Although upwards of 40,000 sets of data on azeotropic behavior have been reported, the number of possible azeotropes is theoretically nearly infinite (Kim & Simmrock, 1997). Moreover, the determination of azeotropy strictly from experiment can be expensive and time-consuming. Evidently, the availability of computational methods for predicting azeotropes would not only reduce this cost, but also save time by narrowing the experimental search space.

* To whom correspondence should be addressed. Fax: +55 11 3813 2380
E-mail addresses: rita@lscp.pqi.ep.usp.br; oller@usp.br; quina@usp.br.

Although several practical correlation rules for azeotropic behavior have been proposed, many of these are empirical and lack a clear physical basis. Moreover, their range of applicability is often quite restricted in scope, e.g., it may not be possible to extend them to all types of non-ideal mixtures without recourse to additional assumptions.

With this in mind, the objective of the present work is to apply an alternative approach employing properties of the individual components of the mixture, i.e., without any explicit consideration of non-ideality of mixing of the components in the liquid phase. A set of macroscopic and molecular properties information of the individual components of the system together with available experimental data of azeotropic conditions are correlated using neural networks for predicting whether a given liquid binary mixture should or should not be expected to form an azeotrope and their azeotropic point. The macroscopic properties examined include the molecular weight, boiling point, critical constants and acentric factor, while the molecular properties consist of parameters related to the molecular dipolarity and the propensity of each component to participate in intermolecular hydrogen bonding interactions. The advantage of neural networks to correlate all these information is that they do not require the use of any explicit models, thermodynamic or otherwise, for the non-ideality of mixture of the components in the liquid phase, as cited before.

Initially, a general review of azeotropy and neural networks is presented. Subsequently, the methodology used and structure of the net are outlined, including the best choice of input parameters, as well as the rules established to discriminate whether a given binary mixture does or does not form an azeotrope and the azeotropic conditions.

2. Review

2.1. Azeotropy

The word "azeotrope" is synonymous with "constant-boiling mixture". An understanding of azeotropes is desirable for two reasons: first, they often occur in fractionation and make a given separation impossible by ordinary rectification; secondly, they may be utilized to separate mixtures not ordinarily separable by simple rectification.

The formation of constant-boiling mixtures is a function of the two following factors:

- > Non-ideality of the solution;
- > Difference in boiling point between the two pure components.

In an ideal solution, the average intermolecular forces are unchanged upon mixing; likewise, there are no changes in internal energy or in volume, no heat is absorbed or evolved, and the vapor pressure of the mixture is a linear function of the molar composition, i.e., the system obeys Raoult's Law.

Since the formation of azeotropic mixtures is a function of the non-ideality of mixing of the components, a correlation of the deviations from ideality, i.e., from Raoult's law, for several mixtures would greatly aid in the prediction and selection of azeotrope-forming systems. In general, the deviations from ideality can be attributed to contributions from the following different types of intermolecular forces: (1) dispersion interactions, (2) dipole-dipole interactions, (3) dipole-induced dipole interactions and (4) hydrogen bonding, this last being one of the most important contributors to azeotrope formation.

However, since the formation of binary azeotropic systems also depends on the difference between the boiling points of the two components, this difference must be taken into account for more accurate predictions. The azeotropic range concept introduced by Swietoslowski (1950) and developed by Malensinski (1956) and Yoshimoto & Mashiko (1956) can explain this effect of boiling point difference on azeotropy.

The more ideal the binary mixture is, the less likely it will be to form an azeotropic system. The closer the boiling points of the two components, the more likely they will form an azeotrope. Therefore, for any given pair of substances, the smaller the deviation from

ideality, the smaller the difference in boiling points required for azeotrope formation. Conversely, components with widely different boiling points may not form an azeotrope even though their binary mixture exhibits very non-ideal behavior.

From a practical standpoint, a prediction method based on the difference in boiling points of the components and their propensity for forming a non-ideal mixture would, therefore, appear to be the best candidate for correlating azeotrope formation with physical and molecular properties of the components.

2.2. Neural Networks

In the context of chemical engineering applications, neural networks are of particular interest as predictive models (Nascimento et al., 1998, 1999 and 2000) and for pattern recognition (Nascimento et al., 1997). Moreover, neural networks have become a widely used tool in areas such as process analysis, simulation, and control, primarily due to their inherent advantages such as e.g. adaptability, non-linearity, and fault tolerance especially in combination with the relatively simple way of use. Neural networks possess the ability to 'learn' the behavior of the process without actually specifying the physical and chemical laws that govern the system. In addition, neural network models can simultaneously employ continuous and discrete input variables (Migliavacca et al., 1999). The success in obtaining a reliable and robust network strongly depends on the choice of appropriate input or process variables, as well as the available set of data and the domain used for training purposes.

The neural network (NN) employed in this study is a three-layer feed-forward network. This type of network, in which information propagates in only one direction, is particularly useful for steady-state modeling. The architecture of a typical NN is shown in Figure 1. The nodes in different layers of the network represent "neuron-like" processing elements. The number of neurons in input and output layers depends on the respective number of inputs and outputs being considered. On the other hand, hidden layers may vary, depending on the user specification, and its structure defines the feed forward NN topology. The neurons in the input layer do not perform data processing functions; they merely provide a means by which scaled data is introduced into the network. These signals are then "fed forward" through the net via the connections, through hidden layers, and eventually to the final output layer. Each interconnection has associated with it a weight that modifies the strength of the signal flowing along that path. Thus, the input to each neuron is a weighted sum of the outputs from neurons in the previous layer. The output of each node is obtained by passing the weighted sum through a nonlinear operator, in this work, a sigmoid function, which can be considered as an universal approximator and is able to fit convex and concave multivariable functions (Cybenko, 1989). This function provides the network with the ability to represent a nonlinear relationship. Selection of an optimal network structure is crucial for network performance. For example, a network with too few hidden nodes cannot represent the system satisfactorily and a network with too many hidden nodes overlearns particular runs presented in training set.

Once the network topology is specified, a set of input-output data is used to train the network; i.e., to determine appropriate values for the weights associated with each interconnection. The data are propagated forward through the network to produce an output that is compared with the corresponding output in data set, hence generating an error. This error is minimized by making changes to the weights, and may involve many passes through training data set. For a given topology, the magnitude of the weights defines the network characteristics and the structural properties of the model. Thus, a NN has the capability to represent complex systems whose structural properties are unknown (Willis et al., 1991).

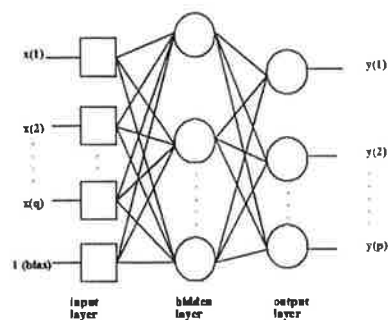


Fig. 1. Structure of a feed forward neural network

The typical way of network design consists of training various network structures with one part of the data, the so-called training set, and their performance is evaluated afterwards with the remaining part of the data, the so-called test set. Both sets can be made large enough to ensure satisfactory learning of the network weights and accurate evaluation of the network performance (Schenker and Agarwal, 1996). The algorithm adopted for the learning phase is the back-propagation algorithm, which is a generalization of the steepest descent method (Haykin, 1994).

There are a large number of parameters (weights) involved in the characterization of neural networks. These parameters reflect the high connectivity among the neurons. Normally, the data are split into two sets: the learning-set, which is used to train the network, i.e., to fit the data and should preferably contain the extremes of all the input and output variables; and the test, which is used to verify the prediction capability of the trained network. In training a network, the objective is to find an optimum set of weights. When the number of weights is larger than the number of available data points, the error in fitting the non-trained data initially decreases, but then increases as the network becomes over trained. In contrast, when the number of weights is much smaller than the number of data points, the overfitting problem is not crucial.

3. Methodology

The basic model structure consists in two distinct NN. The first one (NN1) correlates the properties of pure compounds in order to predict whether or not a given binary mixture should be expected to form an azeotrope. If the azeotrope exists, then a second NN (NN2) correlates the properties of pure compounds of the binary systems with the azeotropic point (type, temperature and composition). The complete model is shown schematically in Figure 2.

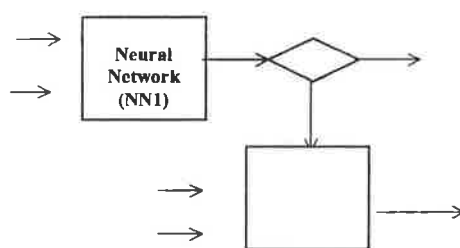


Fig.2. Neural Network Scheme for azeotropy prediction

The methodology employed in this work consists of the following three steps:

- Prediction whether or not a given binary mixture is an azeotropic mixture
- Prediction of the type of azeotrope
- Prediction of the azeotropic conditions

3.1. Prediction whether or not a given binary mixture is an azeotropic mixture

This step is described in details below:

Choice of the appropriate information variables

This step involves the selection of a reliable set of pure substance properties relevant to the characterization of binary mixtures with respect to azeotrope formation. In general, the methods that have been proposed for the prediction and correlation of azeotropy make use of thermodynamic models. Alternative methods are based on a consideration of the intermolecular forces that can potentially contribute to non-ideality. Thus, for example, one can attempt to predict the degree of non-ideality of mixing from the hydrogen bonding capabilities of certain functional groups present in the compounds or from data for homologous series, without taking into account other intermolecular forces. Although the predictions provided by such methods often show reasonable accuracy, the use of these methods is limited to fairly similar types of compounds.

For discussion of various kinds of azeotropic mixtures, it has chosen to adopt the classification proposed by Yoshimoto and Mashiko (1956), which groups binary azeotropic mixtures into four classes corresponding to the predominant type of intermolecular force involved in the mixing.

The problem now is to identify a set of pure substance properties capable of representing these phenomena. A priori, one should consider examining both macroscopic and molecular properties of the individual components of the mixtures as potential variables related to azeotrope formation. These individual single-component physical properties could be molecular weight (MW), boiling point (T_{bp}), critical constants (P_c, T_c e Z_c) and the acentric factor (w). In addition, one requires molecular parameters that reflect the capability to participate in intermolecular interactions, such as the Hildebrand or solubility parameter (δ) and the medium - independent parameters solvatochromic for solute dipolarity (π_2) and solute hydrogen bond acidity ($\Sigma\alpha_2$) and basicity ($\Sigma\beta_2$) developed by Abraham (1993). The acentric factor, ω , takes into account the degree of non-sphericity of the molecule. Differences in shape affect molecular packing and hence the strength of the resultant intermolecular interactions. Indeed, azeotropy is more prevalent when the component molecules are non-spherically shaped (Yoshimoto & Mashiko, 1956). Berg et al. (1944) classified liquids into five groups according to their hydrogen bond capabilities. Consideration of the groups to which each of the liquids belongs provides a basis for qualitative prediction of their potential for hydrogen bond formation and hence of the nature of the deviations from ideality and the likelihood of azeotrope formation. The dipolarity parameter π_2 reflects the propensity to participate in both dipole-dipole and molecular polarizability-dependent dipole-induced dipole interactions.

The basic macroscopic properties were obtained from Reid et al. (1986), the solubility parameters from the DIPPR data bank and the solvatochromic parameters from literature compilations (Abraham, 1993; Abraham et al., 1994). The evaluation of these parameters with respect to their importance for azeotropy prediction is described below in the section on neural network selection and training.

Azeotropic Database

The experimental data included in the binary azeotropic database were taken largely from the compilations of Horsley (1973) and Gmehling et al. (1994). Careful attention was paid to the evaluation of the accuracy of the data regarding azeotrope formation and the

data reported by different authors were checked before inclusion in the database. Unfortunately, in some cases literature contains conflicting reports such as the existence or not of azeotropes. In these cases, the most recently published study was generally considered to be the most reliable. The final database contains about 500 binary azeotropic data sets on the basis of cluster analysis to insure similar distributions over the multivariable domain within each data set.

Neural Network Selection and Training

A three-layered feed-forward neural network was used and trained with the back-propagation algorithm. As noted above, the success in obtaining a reliable and robust network depends strongly on the choice of the variables involved, as well as the nature and domain of data used for training purposes (Chan & Nascimento, 1994)

In this work, the neural network input variables were the pure substance properties or molecular parameters and the output variable the azeotropic information. These input and output variables constitute a given data set. Thus, each data set is a vector containing information about two pure substances and their capacity to form (or not form) an azeotropic binary mixture. The information is arranged in the following manner: first, the properties/parameters of the more volatile component, followed by those of the less volatile component (in the same order) and, finally, either a value of 1 for an azeotrope forming system or 0 for a non-azeotrope-forming mixture.

The experimental azeotropic data were divided into three groups: a "learning set" (LEARN), a "test set" (TEST) and a "validation set" (VALID). LEARN was used to train the neural network, i.e., to obtain the weights of the neural network system. TEST was used to identify the best set of weights; while the third set, VALID, was employed as a completely independent set to validate the neural network. Overfitting was avoided by choosing the neural network that had the lowest possible number of presentations (interactions) and the smallest number of neurons. The primary criterion for selecting the more reliable neural network was a smaller error for the test set. Values of the output variable greater than or equal to 0.8 were used to discriminate an azeotrope forming binary mixture and less than or equal to 0.2 a non-azeotrope forming mixture. Values between these two cutoff limits were classified as cases where no prediction could be made based on the neural network model.

In developing the final neural network model, various sets of input parameters were tested employing network configurations with different numbers of neurons in the hidden layer. The data were assigned to the LEARN and TEST subsets in several different manners and the effect of variation of the number of presentations during neural network training was examined. The best neural network configuration was considered to be that which provided the smallest error in the predictions for the test set.

Briefly, the development of the neural network involved the following sequence of studies. Initially, the input to the neural network during the learning process consisted of ten variables, these being the molecular weight (MW), boiling point (BP), critical temperature (T_c), critical pressure (P_c) and critical compressibility factor (Z_c) for each of the two pure components that constitute the mixture. Data for 65 binary systems were divided into LEARN (47 data sets) and TEST (18 data sets), while data for an additional 26 binary systems was used as the VALID set. Though encouraging, the observed agreement between the experimental and predicted values was only 77%. For a somewhat larger database, consisting of 130 binary systems distributed between LEARN (100) and TEST (30), the learning process was quite satisfactory, but the TEST results showed only 80% agreement between experiment and prediction.

These preliminary results suggested that, in addition to a larger database for the learning process, the development of an accurate neural network model would require the inclusion of additional input parameters and explicit attention to the domain of the variables of the LEARN, TEST and VALID subsets. With this objective, others properties (acentric factor, ω , solubility parameter, δ , and three solvatochromic solute parameters, $\Sigma\alpha_2$, $\Sigma\beta_2$ and π_2)

were included on the NN1 model. The two first parameters did not improve the solution, and then the final number of input parameters was increased to 16 by inclusion of the three solvatochromic solute parameters for each of the two pure components of the mixture. In addition, the database was expanded to include 490 binary systems for which all input parameters were available. These were grouped into LEARN (405 systems), TEST (43 systems) and VALID (42 systems) subsets on the basis of cluster analysis to insure similar distributions over the multivariable domain within each subset. Under these conditions, a network with 11 neurons in the hidden layer provided the smallest error, with an agreement of 90.7% between the experimental and calculated values for the TEST set, as shown in Figure 3.

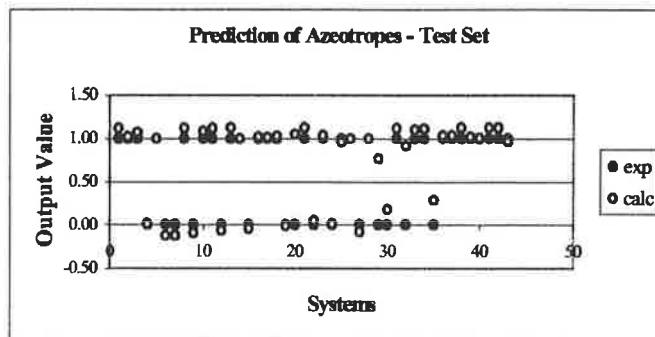


Fig. 3. Prediction for the azeotrope and non-azeotrope binary mixtures - TEST set

Cross correlation and principal component analysis were then used to eliminate variables that were non unique or unrelated to azeotropy. This methodology resulted in the identification of the same input variables as described above.

Model Validation and Analysis of Results

After obtaining the trained neural network, with the appropriate weights that minimize the root mean square error, the network performance was verified employing the validation subset, VALID. The results were very satisfactory, with 88% prediction accuracy. Table 1 shows a comparison between predicted and experimental azeotropy data for the VALID subset utilizing the neural network presented in this paper and the UNIFAC method. Examination of Table 1 reveals that, of the 42 binary mixtures in the VALID set, the neural network model predicted the wrong result in only two cases (ethyl acetate/2-butanone and acetone/benzene) and was unable to make a prediction in three systems (n-hexane/2-pentanone, methyl alcohol/2-pentanone and ethyl alcohol/3-pentanone).

It should be noted that, in all three systems where the neural network fails to make a prediction, one of the components is either 2- or 3-pentanone. A reasonable explanation for this failure is indicated by an analysis of the data from the LEARN sub set. According by Swietoslowski (1959), an azeotropic series is formed by a common component with successive members of a series of homologues, isomers or related substances. In general, because many of the physicochemical properties of homologues of a given series vary in an ordered manner, it can be expected that the azeotropic parameters of such a series should exhibit some regularity too. These characteristics of homologues have indeed often been used to predict the existence of azeotropes (Kim & Simmrock, 1997). However, as shown in Table 2, the binary mixtures of the pentanones with aliphatic alcohols show clear peculiarities in their azeotropic behavior relative to that found for the same alcohols with 2-butanone.

Table 1. Experimental Data and Neural-Network and UNIFAC Predictions Azeotropy for the VALID Set of Binary Mixtures (NN1).

Component 1	Component 2	Exp. Value	NN Method	UNIFAC
n-Pentane	Ethyl Alcohol	Az	Az	Az
Cyclohexane	Butyl Alcohol	Az	Az	Az
Ethyl Alcohol	Chlorobenzene	Non-az	Non-az	Az
Ethyl Acetate	2-Butanone	Az	Non-az	Az
2-Butanone	Methyl Cyclohexane	Az	Az	Az
Ethyl Alcohol	n-Propyl Alcohol	Non-az	Non-az	Non-az
Methyl Alcohol	m-Xylene	Az	Az	Non-az
Acetone	Butyl Alcohol	Non-az	Non-az	Non-az
Acetone	Benzene	Non-az	Az	Non-az
Acetone	1,3-Cyclohexadiene	Az	Az	Az
Ethyl Acetate	Allyl Alcohol	Non-az	Non-az	Non-az
Allyl Alcohol	n-Heptane	Az	Az	Az
n-Propyl Alcohol	sec-Butyl Alcohol	Non-az	Non-az	Non-az
n-Propyl Alcohol	Isobutyl Alcohol	Non-az	Non-az	Non-az
n-Propyl Alcohol	p-Xylene	Az	Az	Az
2-Butanone	Fluor-Benzene	Az	Az	Az
Methyl Cyclopentane	Isobutyl Alcohol	Az	Az	Az
Isoprene	Ethyl Alcohol	Az	Az	Az
Isoprene	Methyl Alcohol	Az	Az	Az
n-Hexane	2-Pentanone	Non-az	NP	Az
Ethyl Alcohol	Butyl Alcohol	Non-az	Non-az	Non-az
Isopropyl Alcohol	Allyl Alcohol	Non-az	Non-az	Non-az
Cyclopentane	Methyl Acetate	Az	Az	Az
Methyl Acetate	Cyclohexane	Az	Az	Az
2-Butanone	sec-Butyl Alcohol	Non-az	Non-az	Non-az
n-Pentane	Benzene	Non-az	Non-az	Non-az
n-Pentane	Cyclohexane	Non-az	Non-az	Non-az
Benzene	Methyl Cyclohexane	Non-az	Non-az	Non-az
n-Hexane	Toluene	Non-az	Non-az	Non-az
n-Heptane	Ethyl Benzene	Non-az	Non-az	Non-az
Toluene	p-Xylene	Non-az	Non-az	Non-az
Methyl Alcohol	Isobutyl Alcohol	Non-az	Non-az	Non-az
Methyl Alcohol	2-Pentanone	Non-az	NP	Non-az
Ethyl Alcohol	3-Pentanone	Az	NP	Az
Allyl Alcohol	Ethyl Benzene	Az	Az	Non-az
Acetonitrile	m-Xylene	Non-az	Non-az	Non-az
Methyl Acetate	n-Propyl Alcohol	Non-az	Non-az	Non-az
Methyl Acetate	Toluene	Non-az	Non-az	Non-az
Isopropyl Alcohol	Butyl Alcohol	Non-az	Non-az	Non-az
2-Butanone	Chlorobenzene	Non-az	Non-az	Non-az
Ethyl Acetate	3-Pentanone	Non-az	Non-az	Non-az
sec-Butyl Alcohol	Isobutyl Alcohol	Non-az	Non-az	Non-az

Notes: Non-az – non-azeotrope-forming system
 Az – Azeotrope-forming systems
 NP – non-predictable (intermediate output value)

Table 2. Experimental azeotropic data for 2-Butanone and 2- and 3-Pentanone.

	2-Butanone	2-Pentanone	3-Pentanone
Methyl Alcohol	Az	Non-az	Non-az
Ethyl Alcohol	Az	Az	Az
n-Propyl Alcohol	Non-az	Az	Az
Isopropyl Alcohol	Az	Non-az	Non-az
Butyl Alcohol	Non-az	Non-az	Non-az
Isobutyl Alcohol	Non-az	Az	Az
sec-Butyl Alcohol	Non-az	-	Az
tert-Butyl Alcohol	Az	-	Non-az

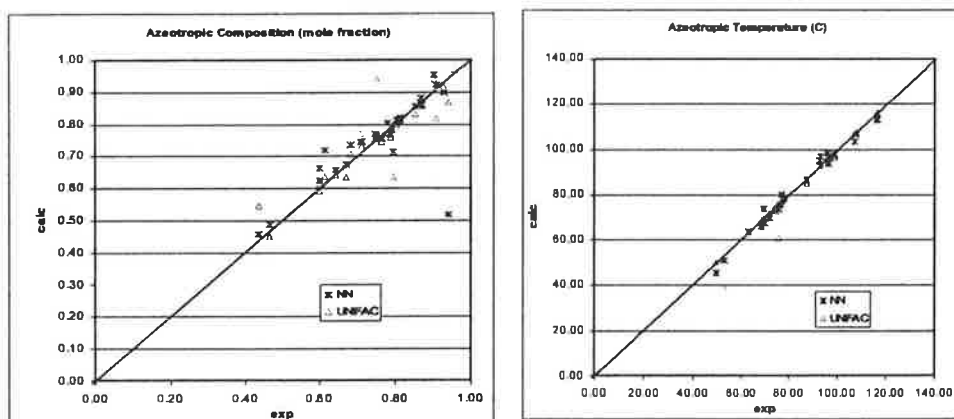
Notes: Non-az – non-azeotrope-forming system Az – Azeotrope-forming systems

Although the UNIFAC method does provide the correct prediction in this case, any model like the neural network that relies on pattern recognition will necessarily fail to predict azeotropy correctly for binary systems of this type.

3.2. Prediction of Type and Azeotropic Conditions

The NN2 to model the prediction of the azeotropic conditions used the same input information (the Abraham parameters and the macroscopic properties) as NN1 used to discriminate whether or not the existence of azeotrope. The NN2 output is the minimum or maximum and homogeneous or heterogeneous azeotrope, coded by integer numbers and the azeotropic conditions (temperature and composition at atmospheric pressure).

The learn set was composed by information of 185 binary system and the validation set was composed by 27 binary system (all these sets are those identified as azeotrope in NN1). No difficulty convergence of the NN2 during the learning procedure was observed. The discrimination of the type of azeotrope, using this NN2, with 11 neurons in the hidden layer, was of 100%. The agreement between the experimental and the predicted values were very good. Figures 4a and 4b show the results. It can be observed that the azeotropic temperature was very good represented and for the azeotropic composition only one point was not discriminated while the others were satisfactorily represented. These figures show also the comparison against predicted values by UNIFAC method. In general, it appears that NN model can predict better the cases presented.



Figs. 4a/b. Prediction of Azeotropic Conditions by Neural Network – TEST SET

4. Conclusions

The prediction of azeotrope formation in binary mixtures using only macroscopic and microscopic properties of the pure components, with no explicit consideration of the non-ideality of mixture, is a very interesting challenge for neural network methodology. The present work demonstrates that a relatively straightforward neural network model with physically reasonable input variables for the pure components can provide a quite good correlation for predictions of the occurrence (or absence) of azeotropy as well as their azeotropic conditions in a wide variety of binary systems. In a relatively small number of situations, the model fails to make a clear prediction regarding azeotropy. All of these correspond predominantly to cases in which structurally homologous molecules exhibit quite distinct azeotropic behavior. Comparison between NN model and UNIFAC method

shows that NN model could predict the azeotropic conditions better than UNIFAC, but conclusive results need more evaluation with other different binary systems.

Acknowledgment

The authors would like to thank FAPESP for financial support of this work.

References

- Abraham, M. H. (1993). *Chem. Soc. Rev.*, 22, 73.
- Abraham, M. H.; Chadham, H. S.; Whiting, G. S.; Michell, R. C. (1994). *J. Pharm. Sci.*, 83, 1085.
- Berg, L., Ewell, R.H. & Harrison, J.M. (1944). Azeotropic Distillation, *Chem. Engng. Progr.*, 36 (10), 871.
- Chan, W.M. & Nascimento, C.A.O. (1994). "Use of Neural Networks for Modeling of Olefin Polymerization in High Pressure Tubular Reactors", *J. Applied Polymer Science*, 53, 1277-1289.
- Cybenko, G., 1989. "Approximation by Superpositions of a Sigmoidal Function", *Mathematics of Control, Signals, and Systems*, vol.2, 303-314.
- Gmehling, J., Menke, J., Krafczyk, J. & Fischer, K. (1994). *Azeotropic Data*, 2 vol., CH-Verlag, Weinheim.
- Haykin, S. (1994). *Neural Networks – A Comprehensive Foundation*. New York:Macmillan College Publishing Company.
- Horsley, L.H. (1973). *Azeotropic Data – III*. Advances in Chemistry Series, No. 116. American Chemical Society, Washington DC.
- Kim, Y.J. & Simmrock, K.H. (1997). AZEOPERT: An Expert System for the Prediction of Azeotrope Formation – I. Binary Azeotropes, *Computers and Chemical Engineering*, 21, 93-111.
- Malesinski, W. (1956). *Bull. Acad. Polon. Sci. Cl.III*, 4, 295, 303, 365, 371.
- Migliavacca, S.C.P., Nascimento, C.A.O. & Rodrigues, C. (1999). "Use of Neural Network for the Simulation of a Gas Centrifuge", *J. Nuclear Science. & Technology*, 36(4), 364-370.
- Nascimento, C.A.O., Guardani, R. & Giuliotti, M. (1997). "Use of Neural Networks in the Analysis of Particle Size Distributions by Laser Diffraction", *Powder Technology*, vol. 90, 89-94.
- Nascimento, C.A.O. & Giudici, G., (1998). "Neural Network Based Approach for Optimization Applied to an Industrial Nylon-6,6 Polymerization Process", *Computers and Chemical Engineering*, vol.22, 595-600.
- Nascimento, C.A.O., Giudici, G. & Scherbakoff, N. (1999). "Modeling of Industrial Nylon-6,6 Polycondensation Process. Part 2: Neural Networks and Hybrid Models.", *J. Applied Polymer Science*, 72, 905-912.
- Nascimento, C.A.O., Giudici, G., Guardani, R. (2000). "Neural Network Based Approach for Optimization of Industrial Chemical Processes", *Computers and Chemical Engineering*, vol.24, 2303-2314.
- Reid, R.C., Prausnitz, J.M. & Poling, B.E. (1988). *The Properties of Gases & Liquids*, McGraw-Hill Int. Ed., 4th ed.
- Schenker, B. and Agarwal, M. (1996). "Cross-Validated Structures Selection for Neural Networks", *Comp. Chem. Engng.*, 20 (2), 175-186.
- Swietoslawski, W. (1950). *Bull. Acad. Polon. Sci.*, 19, 29.
- Yoshimoto, T. & Mashiko, Y. I. (1956). Studies on Azeotropic Mixtures I, *Bull. Chem. Soc. Jpn.*, 29, 990.
- Willis, M.J., Di Massimo, C., Montague, G.A., Tham, M.T. and Morris, A.J. (1991). "Artificial Neural Networks in Process Engineerings". *IEE Proceedings-D*, 138 (3), 256-266.

GROSS ERRORS DETECTION OF INDUSTRIAL DATA BY NEURAL NETWORK AND CLUSTER TECHNIQUES

R.M.B.Alves* and C.A.O.Nascimento

LSCP, Laboratório de Simulação e Controle de Processos, Departamento de Engenharia Química
Escola Politécnica da Universidade de São Paulo,
Av. Prof. Luciano Gualberto, n. 380, trav. 3, 05508-900, São Paulo - SP, Brazil.
E-mail: rita@lscp.pqi.ep.usp.br;
E-mail:oller@usp.br

(Received: March 5, 2002 ; Accepted: June 21, 2002)

Abstract - This article describes the application of a three-layer feed-forward neural network to analyze industrial plant data. To adjust mathematical models (for control or optimization purposes) from plant data, it is necessary to analyze and detect outliers and systematic errors and to remove them. The system studied is the feed preparation of an isoprene production unit and represents a multivariable problem. To detect outliers in a multivariable system is not an easy task. The technique used in this paper is able to identify this kind of error. The methodology employed involves construction of a reliable neural network model to represent the process and its training with a few iterations (a few thousand). Thus, the points at which errors between the experimental and calculated data appear to be scattered far from the majority of the values are probably outliers. In some cases, outlier points can be easily detected, but in others, they are not so obvious. In these cases, they are separated and a cluster with other similar data is built. After analyzing these clusters based on the similarity principle or by hypothesis tests for means, it is then decided whether or not these points can be excluded. At the same time the process is checked for any abnormalities recorded during the specific period. Three year's worth of process data were analyzed and about 30% of the data were excluded.

Keywords: gross error, neural network, modeling, data analysis.

INTRODUCTION

Gross errors or anomalous measurements may arise in the data set due to changes in conditions during the plant operation, errors in the operation of measurement and recording devices, or simply errors in the information register, which may contaminate the valid data. Depending on the average time for data treatment, fluctuations in data could be incorporated in the results. Many times this could result in unreliable information. In cases of errors due to measurement instruments over a long period of time, the average reflects this error. On the other hand, the outlier may simply be one of the extreme values in a probability distribution for a random

variable, which occurs quite naturally but not frequently and should not be rejected.

If one knows the origin of the abnormal values, one does not hesitate to discard this observation. On the other hand, when one is not sure about the error or does not have enough practice to either accept or reject an extreme observation, judgment must be based on some kind of statistical analysis. The question to be analyzed is how probable it is that the observed differences are due solely to random sampling errors so that the decision of whether or not to reject the information can be made. This task becomes especially complicated for complex processes where not all of the influencing parameters are directly accessible or where large stochastic

*To whom correspondence should be addressed

deviations of the process variables result in considerable scattering of the measured data. For this reason, a large variety of approaches to tackling this problem have been proposed. These are commonly based on either statistics or first principle equations or a combination of both. However, this procedure may become extremely complicated either if the underlying physics and chemistry of the process are not very well understood or if application of a sharp statistical criterion for separation of the data into one set of valid and another of invalid values is impossible.

This article demonstrates the ability of a neural network model to learn and adapt itself to different statistical distributions of inputs involving nonlinear mappings. In this way, it allows classification of similar inputs and outputs in order to identify clusters and then proceed with elimination of the gross errors. As will be shown, this approach to detect outliers has considerable potential in the field of data analysis; it is easier and requires much less knowledge of the underlying physicochemical process.

NEURAL NETWORK

Neural computation has become an established discipline and has attracted extensive interest within chemical engineering. Most chemical engineering processes are nonlinear and complex with conventional modeling and simulation techniques often relying on specific simplifying transport, kinetic and/or thermodynamic assumptions. Artificial neural networks (NNs), on the other hand, are able to extract information from a data plant in an efficient manner. NNs have been successfully employed in solving problems in areas such as fault diagnosis, dynamic modeling and control of chemical processes (Bhat and McAvoy, 1990; Hoskins and Himmelblau, 1988; Giudici et al., 1999) and in solving nonlinear optimization problems (Nascimento & Giudici, 1998, Nascimento et al., 2000), among others. In spite of the extensive range of NN applications explored for use in chemical engineering, the quality of information is crucial to train the net and also to avoid overfitting.

Artificial neural networks are made up of highly interconnected layers of simple neuron like nodes. The neurons act as nonlinear processing elements within the network. An attractive property of artificial neural networks is that, given the appropriate network topology, they are capable of characterizing nonlinear functional relationships, representing internal models of a system through a direct learning algorithm, and thus they are able to handle the intrinsic complexities of chemical

processes. Of the many existing artificial neural network paradigms, the three-layers feed-forward neural network is the most widely used network for chemical engineering applications. This NN is classified as a supervised learning network, in which knowledge is captured by the strength of its interconnections between a set of artificial neurons. These interconnections are called the weights of the neural model, which are calculated iteratively using a backpropagation algorithm, i.e., the steepest descent-based optimization routine in order to minimize a given objective function (Rumelhart & McClelland, 1986). The computations are carried out over the entire network, except the input layer. The mapping of each unit is in terms of the combination of all its inputs, followed by the application of a nonlinear function, called the activation function. In this work, a sigmoid function was used as the activation function.

Construction and training of the NN used in this work were carried out using in-house software.

METHODOLOGY

The available industrial data on the process studied (isoprene production unit) were provided as a daily average. These data were collected during three years. Analysis was carried out for each year individually. Treatment of the data was performed in two steps: first, a preliminary analysis for abnormal values (e.g., points out side of the possible operational range, which may be subject to rejection) was carried out. The second step involved data analysis using neural network approach and statistical techniques. The methodology applied in this work follows the steps shown in Figure 1.

The first step in the data analysis makes use of the following criteria to eliminate abnormal values: values out side of an acceptable range for the corresponding variable, graphic observations of the variables as function of time, experience with statistical features as much as with the process and material and energy balances. The variables of interest were defined considering the available process data and its importance to the process and plant operation. Then, the minimum, maximum and mean values as well as the variance for each selected variable were identified. The variables whose operational ranges were too close to the wind-up measurement instruments were not included as neural network information.

After this initial analysis, the resulting data set was used to train the neural network, as shown in Figure 1. In this step, all data were included in the training data set and during the training process few iterations were performed. It was observed that some

points and groups of points were not well adjusted. Identification of these groups (or points) is an indication of consistency problems or gross errors not identified in the first step of the procedure. To decide whether or not these points must be eliminated some statistical analyses such as cluster analysis and hypothesis tests for means were used.

Cluster analysis is based on the similarity principle among several data sets. A data set was formed by the input and output variables chosen for each process unit, corresponding to information from one day of operation. It is expected that for a series of similar input variables, the process must yield similar output variables (dependent variables). When a different input or output variable is observed in a series of similar data, the corresponding data set may be rejected. Table 1 shows two examples of cluster analysis: it can be observed that for variable out2 the values 25.24 and 20.85 in the first and second groups of data, respectively, must be rejected.

In some cases a simple and direct analysis is not possible, e.g., when a given data set is unique or

when there are only two data sets for comparison with some different information, it is not possible to determine which one is correct. In these cases, the domains of the variables are extended, compared with the previous group. Although these new groups are less accurate, it is usually possible to identify abnormal points. For this step, the hypothesis test for means, which involves a confidence interval estimate and a hypothesis test was employed with a confidence level of 95% (Himmelblau, 1970).

Table 2 shows the application of this methodology in analysis of plant data. The values in bold in groups 1 and 2 were not well adjusted during neural network training and it was not possible to identify groups of similar data sets for the cluster analysis; thus a hypothesis test for means analysis was performed. It was observed that the value 25.81 is inside the confidence interval and the null hypothesis is accepted and this data must not be eliminated. On the other hand, the value 1.62 is outside the confidence interval and the null hypothesis is rejected and this data set must be eliminated.

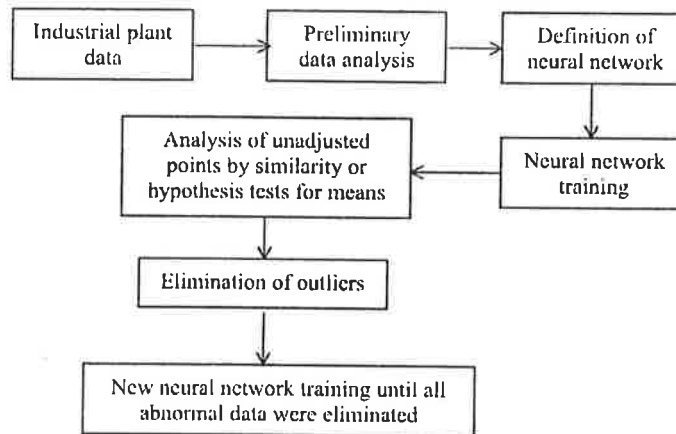


Figure1: Data analysis methodology

Table 1: Cluster analysis - examples

Input Variables						Output Variables					
in1	in2	in3	in4	in5	in6	in7	in8	out1	out2	out3	out4
12.79	15.34	1.86	63.80	39.30	59.90	8.89	4.37	8.35	25.24	2.50	97.4
12.80	15.40	1.88	63.80	39.20	59.90	8.87	4.35	8.35	27.21	2.63	97.3
12.80	15.25	1.79	63.80	39.20	59.90	8.85	4.38	8.35	26.88	2.53	97.4
12.75	15.07	1.71	63.80	39.10	59.80	8.79	4.36	8.35	27.21	2.49	97.3
13.99	16.03	2.40	64.50	55.10	59.80	11.19	4.83	9.08	20.85	2.68	102.6
13.82	16.15	2.93	64.30	55.50	59.30	10.79	5.06	9.16	28.84	2.89	97.2
13.80	16.09	2.91	64.30	55.10	59.40	10.81	4.97	9.23	28.57	2.85	97.8
13.70	16.02	2.74	64.20	55.60	58.60	11.13	4.91	9.07	30.50	2.66	95.0
13.70	15.86	2.54	64.20	55.80	58.60	11.15	5.07	8.92	30.92	2.54	94.6

Table 2: Hypothesis test for means

Group	Input Variables								Output Variables			
	in1	in2	in3	in4	in5	in6	in7	in8	out1	out2	out3	out4
1	13.00	14.50	1.95	66.00	51.10	59.90	10.51	5.55	7.62	26.73	1.62	93.5
	13.00	14.57	2.77	65.90	50.60	60.00	10.25	5.37	7.72	26.15	2.68	91.6
2	13.20	16.20	2.43	64.30	50.40	59.70	9.94	4.41	8.68	25.81	3.17	100.1
Sample	12.00	13.58	2.38	66.00	50.20	61.20	10.40	3.77	8.78	19.91	2.77	97.2
	12.00	14.74	2.14	65.90	48.40	61.30	10.08	3.77	8.89	20.77	2.70	98.5
	12.09	14.13	1.75	66.50	48.60	61.00	10.23	4.04	8.58	23.51	2.07	97.5
	13.30	14.99	2.46	66.80	48.80	60.90	10.32	5.15	8.73	24.35	2.70	94.8
	13.48	15.74	2.37	66.80	49.60	60.70	10.47	5.35	8.76	24.60	2.47	93.8
	13.50	15.64	2.23	66.80	50.30	60.80	10.57	5.48	8.72	25.22	2.30	93.8
	13.51	15.02	2.01	66.70	50.50	60.80	10.56	5.51	8.76	25.23	2.26	94.0
	13.50	15.87	1.96	66.40	49.10	60.60	10.34	5.49	8.79	25.91	2.17	94.5
	13.00	14.50	1.95	66.00	51.10	59.90	10.51	5.55	7.62	26.73	1.62	93.5
	12.98	13.52	2.22	65.70	51.60	60.00	10.65	5.38	7.74	26.09	2.57	93.1
	13.00	14.57	2.77	65.90	50.60	60.00	10.25	5.37	7.72	26.15	2.68	91.6
	13.20	16.20	2.43	64.30	50.40	59.70	9.94	4.41	8.68	25.81	3.17	100.1
	13.00	14.06	2.05	64.30	51.20	59.70	10.27	4.28	8.65	20.93	2.66	100.5
	13.38	14.13	1.98	64.30	50.10	59.40	10.38	4.41	8.84	20.98	2.52	100.8
	14.01	13.99	2.60	64.50	52.30	59.60	10.97	4.40	9.45	17.98	2.94	100.9
	14.00	14.24	2.64	64.40	52.30	59.50	11.00	4.61	9.30	18.49	2.89	99.6
	14.00	14.11	2.57	64.40	52.30	59.50	11.00	4.62	9.30	19.09	2.89	99.9
	14.01	15.67	2.59	64.30	50.70	59.70	10.81	4.56	9.41	19.65	2.82	99.8
Minimum	12.00	13.52	1.75	64.30	48.40	59.40	9.94	3.77	7.62	17.98	1.62	91.60
Maximum	14.01	16.20	2.77	66.80	52.30	61.30	11.00	5.55	9.45	26.73	3.17	100.90
Mean	13.22	14.71	2.28	65.56	50.45	60.24	10.49	4.79	8.71	22.86	2.57	96.88
Std dev	0.65	0.82	0.29	1.04	1.23	0.66	0.31	0.63	0.54	3.06	0.37	3.20
t(n-1=18) 2.11												
Mean+std dev*t	14.60	16.44	2.90	67.75	53.05	61.63	11.13	6.10	9.84	29.31	3.35	103.63
Mean-std dev*t	11.84	12.97	1.67	63.37	47.85	58.85	9.84	3.47	7.57	16.41	1.78	90.13

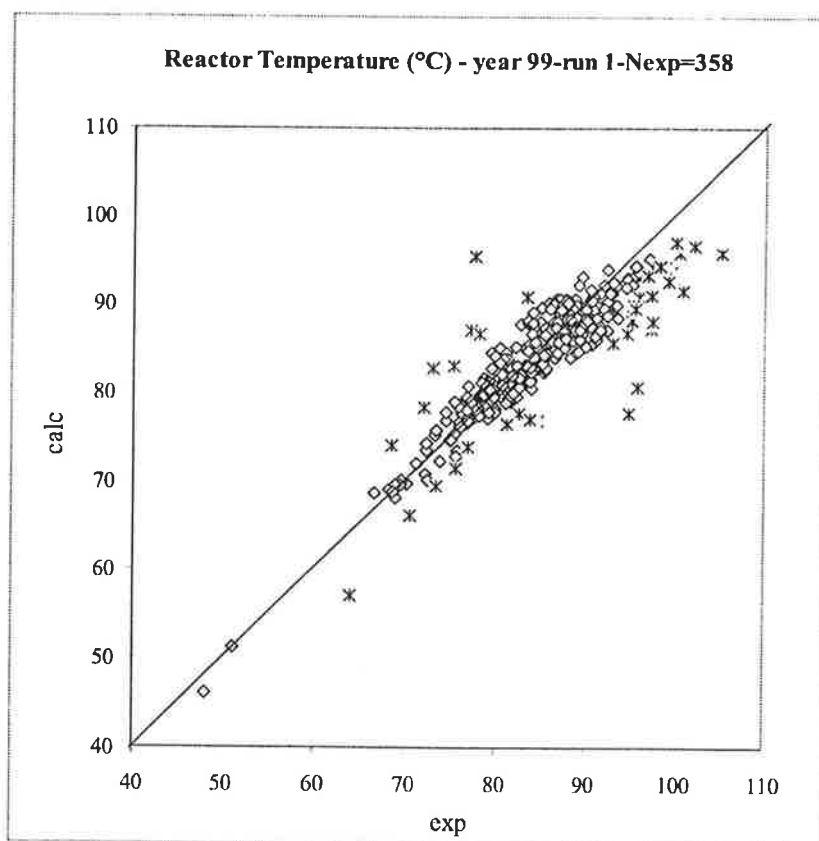
RESULTS AND DISCUSSION

A simple measure to assess the quality of the fit of the chosen neural network to the experimental data is usually a comparison of the values calculated by the neural network with the original experimental data. The scatter of data points around the ideal 45° line can be used to judge the fit of the neural network to the experimental data. The idea to use neural networks for the purpose of outlier detection is based on this kind of diagram (Bülau et al., 1999). Hence, it must only be shown that the outliers of the experimental data correspond to the outliers from this curve. Thus, the

neural network was first of all trained for the entire data set and afterwards for the filtered data set. The outliers detected after the first training run were analyzed by application of the statistics described previously. This procedure was repeated several times until the scattered data no longer showed abnormal points. Since the training of the network with the filtered data leads to results for calculated data, which are different from those for the original data set, the input database changes due to the filtration procedure. Table 3 shows the number of eliminated points in each run and Figures 2a-c show the results of this method for the first, second and final runs, respectively.

Table 3: Points eliminated in each run

Run	Number of points eliminated
1	49
2	28
3	19
4	7
5	12
6	-



(a)

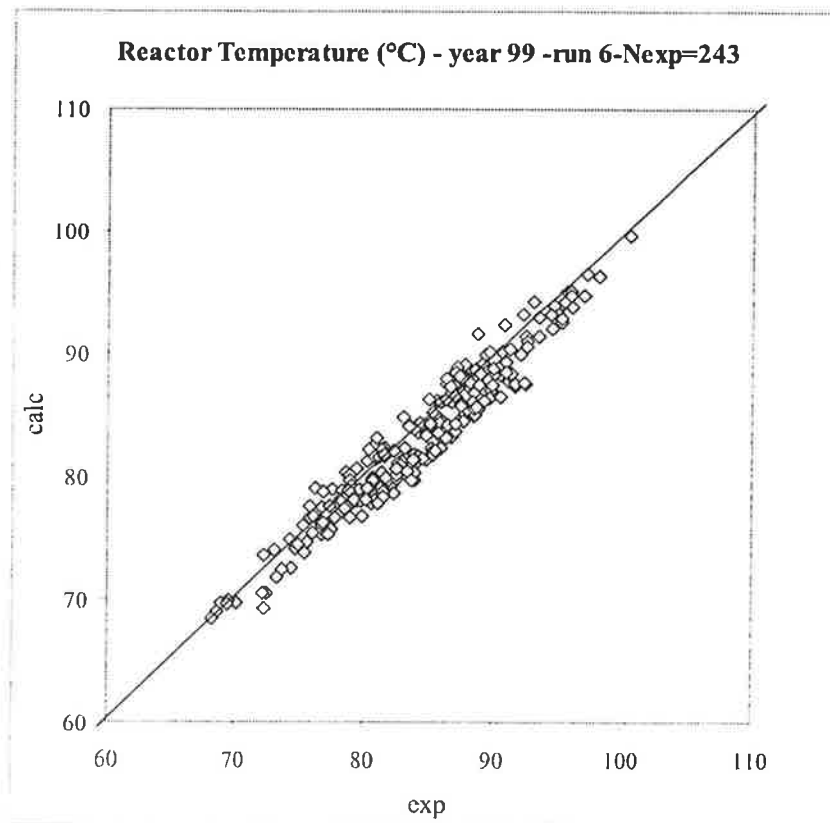
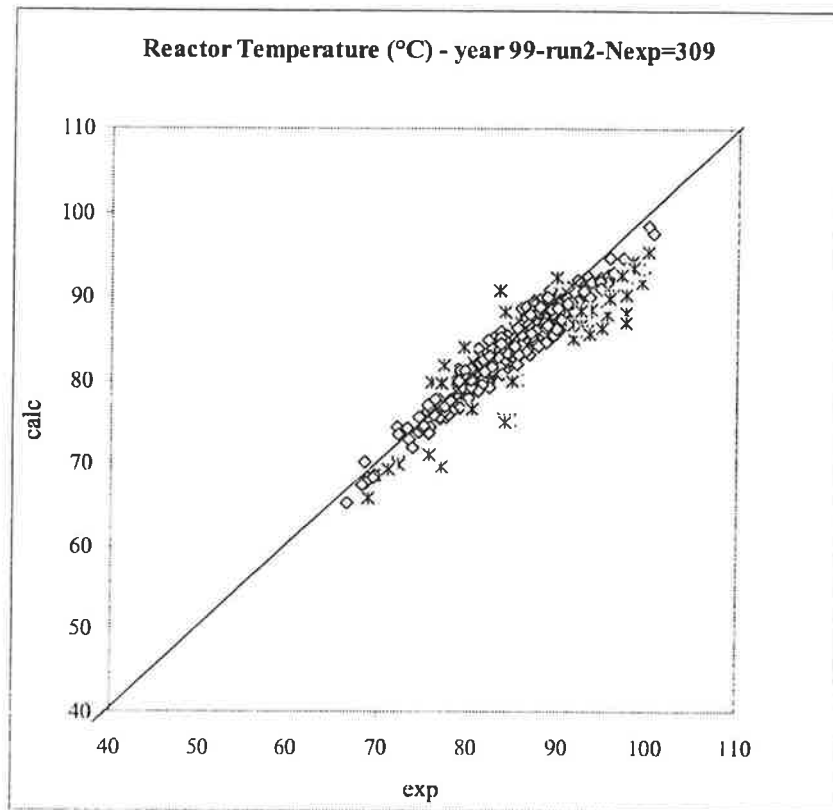


Figure 2: Comparison of calculated and measured data: (a) before analysis; (b) intermediate results; (c) final results

CONCLUSIONS

Application of a neural network is a very attractive tool for detecting outliers. It is simple, more cost effective and more easily used, particularly by plant engineers, and the results presented demonstrate that neural networks have considerable potential in the field of data analysis, mainly because they require much less knowledge of the underlying physicochemical process. However, the final decision to eliminate the suspect data is made by applying a cluster analysis to this approach to detect outliers.

ACKNOWLEDGMENTS

The authors gratefully acknowledge FAPESP for its financial support and COPENE Petroquímica do Nordeste for providing the industrial data used in this work.

REFERENCES

- Bhat, N. and McAvoy, T.J. (1990). Use of Neural Nets for Dynamic Modeling and Control of Chemical Process Systems. *Comput. Chem. Eng.*, 14 (4/5), 575.
- Büla, H.C., Ulrich, J., Guardani, R. and Nascimento, C.A.O. (1999). Application of Neural Networks to Data from a Melt Crystallization Process for the Detection of Outliers. In Proceedings of AIDA, International Seminar on Advances in Data Analysis - Washington, DC.
- Giudici, R., Nascimento, C.A.O., Tresmondi, A., Domingues, A. and Pellicciotta, R. (1999). Mathematical Modeling of an Industrial Process of Nylon-6,6 Polymerization in a Two-Phase Flow Tubular Reactor. *Chem. Eng. Sci.*, 54, 3242.
- Himmelblau, D.M. (1970). *Process Analysis by Statistical Methods*. John Wiley & Sons, Inc.
- Hoskins, J.C. and Himmelblau, D.M. (1988). Artificial Neural Network Models for Knowledge Representation in Chemical Engineering. *Comput. Chem. Eng.*, 12 (9/10), 881.
- Nascimento, C.A.O. and Giudici, R. (1998). Neural Network Based Approach for Optimization Applied to an Industrial Nylon-6,6 Polymerization Process. *Comput. Chem. Eng.*, 22, S595.
- Nascimento, C.A.O., Giudici, R. and Guardani, R. (2000). Neural Network Based Approach for Optimization of Industrial Chemical Processes. *Comput. Chem. Eng.*, 24, 2303.
- Rumelhart, D.E. and McClelland, J.L. (1986). *Parallel Distributed Processing (Vol. 1)*, Cambridge, MA: MIT Press.

Claudio

From: "Carolyn Blue" <carolyn@ecn.purdue.edu>
To: <oller@usp.br>
Sent: Friday, May 30, 2003 11:03 AM
Subject: ms#2654

Re: MS#2654 - A New Approach for the Prediction of Azeotropy in Binary Systems

Dear Prof. Nascimento:

While updating my database, I noticed that your revisions for the above manuscript were received in January, however, I had been waiting instructions from Prof. Ponton on whether it is now ready for publication. I sent him email and he has informed me to go ahead and accept. So your manuscript has been sent to Elsevier today for publication.

Sorry about the long wait,

Regards,

Carolyn Blue
Editorial Asst.

New Approach for the Prediction of Azeotropy in Binary Systems

Rita Maria Brito Alves¹, Frank H. Quina² and Claudio Augusto Oller Nascimento^{1*}

¹LSCP- Laboratório de Simulação e Controle de Processos, Departamento de Engenharia Química, Escola Politécnica da Universidade de São Paulo, Av. Prof. Luciano Gualberto, n. 380, trav. 3, 05508-900, São Paulo – SP, Brazil

²Instituto de Química, Universidade de São Paulo, CP 26077, São Paulo 05513-970, Brazil

*To whom correspondence should be addressed. Fax:+55-11-38132380
E-mail address: oller@usp.br

ABSTRACT

A new approach for the prediction of azeotrope formation between components in a mixture, that does not require vapor-liquid equilibrium calculations, is presented. The method employs neural networks to correlate azeotropic data for binary mixtures with a series of macroscopic and microscopic properties of the pure components, without explicit consideration of non-ideality of mixture. The model fails to make a clear prediction regarding azeotropy in only a relatively small number of situations in which structurally homologous molecules are known to exhibit quite distinct azeotropic behavior.

Keywords: Neural Network; Azeotropy; Azeotropic prediction

INTRODUCTION

The ability to predict whether a given mixture will form an azeotrope is essential in any separation process. Moreover, an understanding of azeotropy is important for the successful selection of entrainers in azeotropic distillation, of auxiliary materials in extractive distillation and of the appropriate solvents in extractions. Although upwards of 40,000 sets of data on azeotropic behavior have been reported (Horsley, 1973, Gmehling et al., 1994), the number of possible azeotropes is theoretically nearly infinite. Since the determination of azeotropy strictly from experiment can be expensive and time-consuming, the availability of computational methods for predicting azeotropes would not only reduce this cost, but also save time by narrowing the experimental search space. Although several practical correlation rules for azeotropic behavior have been proposed, many of these are empirical, lack a clear physical basis, and are of restricted applicability or require additional assumptions in order to extend them to all types of non-ideal mixtures.

With this in mind, the objective of the present work was to examine the possibility of predicting whether a given binary mixture of two liquids should or should not be expected to form an azeotrope based only on properties of the individual components of the mixture, i.e., without any explicit consideration of non-ideality of mixing of the components in the liquid phase.

PROBLEM FORMULATION AND SOLUTION

The formation of azeotropes, or constant-boiling mixtures, is known to be a function of the non-ideality of the solution, combined with the difference in boiling point between the two pure components. The azeotropic range concept developed by Yoshimoto & Mashiko (1956) can explain this effect of boiling point difference on azeotropy. In general, the deviations from ideality can be attributed to contributions from dispersion interactions, dipole-dipole interactions, dipole-induced dipole interactions and hydrogen bonding, this last being one of the most important contributors to azeotrope formation. These considerations led us to examine both macroscopic and molecular properties of the individual components of the mixture as potential variables related to azeotrope formation. These included single-component physical properties such as the molecular weight (MW), the boiling point (T_{bp}), the critical constants (P_c , T_c , z_c), the Hildebrand solubility parameter (δ) and the acentric factor (ω) and the medium-independent

solvatochromic parameters for solute dipolarity (π_2) and solute hydrogen bond acidity ($\Sigma\alpha_2$) and basicity ($\Sigma\beta_2$) developed by Abraham (1993). The basic macroscopic properties were obtained from Reid et al. (1986), the solubility parameters from the DIPPR data bank and the solvatochromic parameters from literature compilations (Abraham, 1993; Abraham et al., 1994). Cross-correlation and principal component analyses were then employed to eliminate variables that were either non-unique or unrelated to azeotropy. This resulted in the identification of eight unique input variables for each pure component of the azeotrope: T_{bp} , T_c , P_c , π_2 , $\Sigma\alpha_2$, $\Sigma\beta_2$, z_c and MW. Although T_c is known to correlate empirically with T_{bp} and MW (Reid et al., 1988), the correlation has a relatively large error (for the present purposes, at least). Indeed, exclusion of either T_c or MW from the input variable set resulted in neural networks with significantly poorer performance.

In order to correlate these variables with azeotrope formation, we chose to employ neural network methodology. Indeed, in chemical engineering applications, neural networks have proven to be particularly interesting as predictive models (Nascimento et al., 2000). Neural networks possess the ability to 'learn' the behavior of the process without actually specifying the physical and chemical laws that govern the system. In addition, neural network models can simultaneously employ continuous and discrete input variables (Migliavacca et al., 1999). The success in obtaining a reliable and robust network strongly depends on the choice of appropriate input or process variables (*vide supra*), as well as the available set of data and the domain used for training purposes.

The binary azeotrope database, consisting of 490 binary systems for which all input parameters were available, was grouped into "learning" or "training" (405 systems), "test" (43 systems) and "validation" (42 systems) subsets. The data were distributed randomly into three groups and the "test" and the "validation" data were checked to insure that they fell inside the input variable domain of the "training" set. One inevitable consequence of this procedure for our data set was that it attributed all pairs in which one of the components was either water (34 systems) or chloroform (27 systems), all of the heterogeneous azeotropic systems (33, all minimum boiling) and all of the maximum boiling azeotropes (4 systems, all homogeneous) to the training set. A three-layered feed-forward neural network interconnected with conventional sigmoidal transfer functions (Migliavacca et al., 1999) was used throughout. Networks were trained with the back-propagation algorithm. The primary criterion for selecting the more reliable neural network was the smallest error for the test set. For presentation to the network, it is

necessary to order the components in some unambiguous manner. For this purpose, we selected the boiling point, such that, for each mixture, component 1 in the network was always the more volatile liquid and component 2 the less volatile one. During training, azeotrope formation was characterized by a binary indicator variable, the values of unity and zero corresponding to azeotrope-forming and non-azeotrope-forming mixtures, respectively. For validation, values of the corresponding output variable greater than or equal to 0.8 were used to discriminate an azeotrope-forming binary mixture and less than or equal to 0.2 a non-azeotrope-forming mixture. Values between these two cutoff limits were classified as cases where no prediction could be made based on the neural network model. Overfitting was avoided by choosing the neural network that had the lowest possible number of presentations (interactions) and the smallest number of neurons. Under these conditions, a network with 6 neurons in the hidden layer provided the smallest acceptable error, with an agreement of better than 90% between the experimental and calculated values for the test subset, as shown in Figure 1.

RESULTS AND DISCUSSION

After obtaining the trained neural network, with the appropriate weights that minimize the root mean square error, the network performance was verified employing the validation subset. Table 1 compares the experimental azeotropy data for this subset with that predicted by the neural network developed in this paper and by the UNIFAC method.

Of the 42 binary mixtures in this subset, the neural network model predicted the wrong result in only two cases (methanol/2-pentanone and acetone/benzene) and was unable to make a prediction in four systems (1-propanol/2-methyl-1-propanol, n-hexane/2-pentanone, ethyl acetate/2-butanone and ethanol/3-pentanone). In most of the systems where the neural network fails to make a prediction, one of the components is an aliphatic ketone. A reasonable explanation for some of these failures is indicated by the data in Table 2 for the 22 solvent pairs included in the training subset, which show that there are clear peculiarities in the azeotropic behavior of binary mixtures of 2-butanone and the pentanones with aliphatic alcohols. Although the UNIFAC method does provide the correct prediction in this case, any model that relies on pattern recognition, like the neural network approach, will necessarily fail to predict azeotropy adequately for binary systems of this type. In this context, it is important to note that the network, rather than responding incorrectly, makes no prediction with respect to azeotropy.

The use of macroscopic and microscopic properties of the pure components, in particular solvatochromic solute parameters, with no explicit consideration of the non-ideality of mixture, represents a new approach for the prediction of azeotrope formation in binary mixtures. The present work demonstrates that a relatively straightforward neural network model with physically reasonable input variables for the pure components provides a quite good correlation of data for the occurrence (or absence) of azeotropy in a wide variety of binary systems. In a relatively small number of situations, the model does fail to make a clear prediction regarding azeotropy. These correspond predominantly to cases in which structurally homologous molecules exhibit quite distinct azeotropic behavior.

Acknowledgment

The authors thank the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support of this work and fellowships, respectively.

References

- Abraham, M. H. (1993). "Scales of Solute Hydrogen-Bonding - Their Construction and Application to Physicochemical and Biochemical Processes", *Chem. Soc. Rev.*, **22**, 73.
- Abraham, M. H.; Chadham, H. S.; Whiting, G. S.; Michell, R. C. (1994). "Hydrogen-Bonding. 32. An Analysis of Water-Octanol and Water-Alkane Partitioning and the Delta-Log-P Parameter Of Seiler", *J. Pharm. Sci.*, **83**, 1085.
- Gmehling, J., Menke, J., Krafczyk, J. & Fischer, K. (1994). *Azeotropic Data*, 2 vol., CH-Verlag, Weinheim.
- Horsley, L.H. (1973). *Azeotropic Data – III. Advances in Chemistry Series*, No. 116. American Chemical Society, Washington DC.
- Migliavacca, S.C.P., Nascimento, C.A.O. & Rodrigues, C. (1999). "Use of Neural Network for the Simulation of a Gas Centrifuge", *J. Nuclear Science. & Technology*, **36**(4), 364-370.
- Nascimento, C.A.O., Giudici, R. & Guardani, R. (2000). "Neural Network Based Approach for Optimisation of Industrial Chemical Process", *Computers and Chemical Engineering*, **24**, 2303-2314.

Reid, R.C., Prausnitz, J.M. & Poling, B.E. (1988). *The Properties of Gases & Liquids*, McGraw-Hill Int. Ed., 4th ed.

Yoshimoto, T. & Mashiko, Y. I. (1956). Studies on Azeotropic Mixtures I, *Bull. Chem. Soc. Jpn.*, 29, 990.

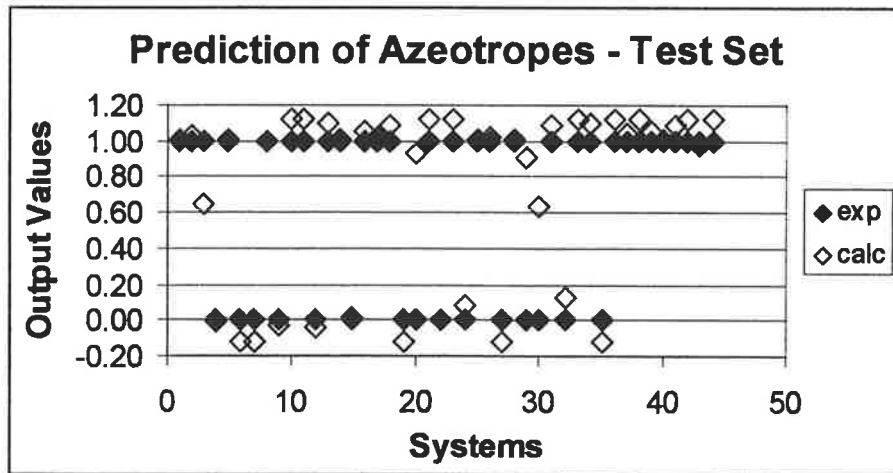


Fig. 1. Prediction for the azeotrope and non-azeotrope-forming binary mixtures of the TEST set

Table 1. Experimental Results and Neural-Network and UNIFAC Predictions of Azeotropy of Binary Mixtures for the Validation Subset.

Component 1	Component 2	Experimental Result	NN Method	UNIFAC Method
n-Pentane	Ethanol	Az	Az	Az
Cyclohexane	1-Butanol	Az	Az	Az
Ethanol	Chlorobenzene	Non-Az	Non-Az	Az
Ethyl Acetate	2-Butanone	Az	NP	Az
2-Butanone	Methyl Cyclohexane	Az	Az	Az
Ethanol	1-Propanol	Non-Az	Non-Az	Non-Az
Methanol	m-Xylene	Az	Az	Non-Az
Acetone	1-Butanol	Non-Az	Non-Az	Non-Az
Acetone	Benzene	Non-Az	Az	Non-Az
Acetone	1,3-Cyclohexadiene	Az	Az	Az
Ethyl Acetate	2-Propen-1-ol	Non-Az	Non-Az	Non-Az
2-Propen-1-ol	n-Heptane	Az	Az	Az
1-Propanol	2-Butanol	Non-Az	Non-Az	Non-Az
1-Propanol	2-Methyl-1-propanol	Non-Az	NP	Non-Az
1-Propanol	p-Xylene	Az	Az	Az
2-Butanone	Fluorobenzene	Az	Az	Az
Methyl Cyclopentane	2-Methyl-1-propanol	Az	Az	Az
Isoprene	Ethanol	Az	Az	Az
Isoprene	Methanol	Az	Az	Az
n-Hexane	2-Pentanone	Non-Az	NP	Az
Ethanol	1-Butanol	Non-Az	Non-Az	Non-Az
2-Propanol	2-Propen-1-ol	Non-Az	Non-Az	Non-Az
Cyclopentane	Methyl Acetate	Az	Az	Az
Methyl Acetate	Cyclohexane	Az	Az	Az
2-Butanone	2-Butanol	Non-Az	Non-Az	Non-Az
n-Pentane	Benzene	Non-Az	Non-Az	Non-Az
n-Pentane	Cyclohexane	Non-Az	Non-Az	Non-Az
Benzene	Methyl Cyclohexane	Non-Az	Non-Az	Non-Az
n-Hexane	Toluene	Non-Az	Non-Az	Non-Az
n-Heptane	Ethylbenzene	Non-Az	Non-Az	Non-Az
Toluene	p-Xylene	Non-Az	Non-Az	Non-Az
Methanol	2-Methyl-1-propanol	Non-Az	Non-Az	Non-Az
Methanol	2-Pentanone	Non-Az	Az	Non-Az
Ethanol	3-Pentanone	Az	NP	Az
2-Propen-1-ol	Ethylbenzene	Az	Az	Non-Az
Acetonitrile	m-Xylene	Non-Az	Non-Az	Non-Az
Methyl Acetate	1-Propanol	Non-Az	Non-Az	Non-Az
Methyl Acetate	Toluene	Non-Az	Non-Az	Non-Az
2-Propanol	1-Butanol	Non-Az	Non-Az	Non-Az
2-Butanone	Chlorobenzene	Non-Az	Non-Az	Non-Az
Ethyl Acetate	3-Pentanone	Non-Az	Non-Az	Non-Az
2-Butanol	2-Methyl-1-propanol	Non-Az	Non-Az	Non-Az

Notes: Non-Az – Non-Azeotrope-forming system
Az – Azeotrope-forming systems
NP – non-predictable (intermediate output value)

Table 2. Experimental azeotropic data for 2-Butanone and 2- and 3-Pentanone.

	2-Butanone	2-Pentanone	3-Pentanone
Methanol	Az	Non-Az	Non-Az
Ethanol	Az	Az	Az
1-Propanol	Non-Az	Az	Az
2-Propanol	Az	Non-Az	Non-Az
1-Butanol	Non-Az	Non-Az	Non-Az
2-Methyl-1-propanol	Non-Az	Az	Az
2-Butanol	Non-Az	-	Az
2-Methyl-2-propanol	Az	-	Non-Az

Notes: Non-Az – Non-Azeotrope-forming system
Az – Azeotrope-forming systems

Appendix: Azeotropy prediction algorithm

The governing equations of the multilayer neural network with one hidden layer like the one used in this work, is given by:

$$S_j = \sum_{i=1}^n w_{i,j} x_i + w_{n+1,j} \quad (\text{a1})$$

where

S_j = input to the j^{th} neuron of the hidden layer;

$w_{i,j}$ = weight between the i^{th} neuron of the input layer and the j^{th} neuron of the hidden layer;

x_i = normalized input variable to the i^{th} neuron.

The outputs from the neurons of the hidden layer (O_j) are calculated by the sigma function $f(S_j)$, given by equation (a3)

$$O_j = f(S_j) \quad (\text{a2})$$

$$f(\alpha) = \frac{1}{1 + e^{-\alpha}} \quad (\text{a3})$$

Similarly, the input to the k^{th} neuron of the output layer is calculated via equation (a4)

$$S_k = \sum_{j=1}^m w_{j,k} O_j + w_{n+1,k} \quad (\text{a4})$$

where

$w_{j,k}$ = weight between the j^{th} neuron of the hidden layer and the k^{th} neuron of the output layer. Since there is only one neuron in the output layer in the present case,

$$y = f(S_k) \quad (\text{a5})$$

The inputs (x_i) and outputs (y) of the network are scaled into the range 0.1 to 0.9. The minimum ($x_{i,\min}$ and y_{\min}) and maximum ($x_{i,\max}$ and y_{\max}) values of each variable are given in Table A

Table A - Values of the minimum and maximum of each variable.

For component 1:

	MW	T_{bp} (K)	T_c (K)	P_c (bar)	Z_c	$\Sigma\alpha_2$	$\Sigma\beta_2$	π_2
$x_{i,\min}$	18.015	169.3	282.4	24.9	0.184	0	0	0
$x_{i,\max}$	119.38	412.3	647.3	221.2	0.310	0.82	0.6	0.9

For component 2:

	MW	T_{bp} (K)	T_c (K)	P_c (bar)	Z_c	$\Sigma\alpha_2$	$\Sigma\beta_2$	π_2
$x_{i,\min}$	18.015	184.6	305.4	24.9	0.184	0	0	0
$x_{i,\max}$	119.38	417.6	647.3	221.2	0.293	0.82	0.6	0.9

The normalization is given by the following equations:

$$x_i^* = 0.1 + \frac{x_i - x_{i,\min}}{\frac{x_{i,\max} - x_{i,\min}}{0.9 - 0.1}} \quad (a6)$$

where

x_i – i^{th} input variable

$x_{i,\min}$ - minimum value of the i^{th} input variable

$x_{i,\max}$ - maximum value of the i^{th} input variable

x_i^* - scaled i^{th} input variable

$$y^* = 0.1 + \frac{(y_{\max} - y_{\min})}{(y_{\max} - y_{\min})} \quad (a7)$$

where

y – output variable

y_{\max} – equals 1 (azeotrope)

y_{\min} – equals 0 (non-azeotrope)

y^* - scaled output variable

The weights ($w_{i,j}$) used in equation (a1) are:

$i \downarrow j \rightarrow$	1	2	3	4	5	6
1	-12.1830	10.0960	16.8020	-13.4700	-3.6154	-10.5380
2	44.8630	-25.0910	-2.7177	-15.6710	-13.6280	61.3580
3	19.2890	-26.0690	-5.5596	-21.6560	-17.1730	21.9800
4	-39.1250	-6.3625	-1.7887	13.5210	-25.8650	32.1190
5	6.5845	-16.0920	17.4850	22.6350	-1.7544	10.0840
6	-30.2190	3.8014	1.6320	5.1146	-10.8540	50.4020
7	-7.9869	24.6020	-15.7850	2.1178	-10.6540	18.9810
8	0.0446	1.3539	-5.1604	-2.9395	-7.8522	23.2210
9	4.4515	-4.6531	-5.8081	9.9132	5.9985	-25.5450
10	-22.8650	-1.1106	-11.1490	56.0010	27.8310	-29.8560
11	-30.3600	4.8724	-4.7142	-8.5689	18.7800	-82.4550
12	20.5050	5.3690	12.2710	-31.7680	2.2473	-16.3800
13	-18.8470	14.8750	-27.0710	-14.4820	-1.4833	18.2200
14	13.8920	-3.0587	-6.3485	5.9634	14.7960	-30.2440
15	16.7140	29.5830	3.0693	-10.7420	-26.6430	-19.4140
16	-6.1907	-22.2140	17.5730	-4.9483	14.6170	-2.1331
17	0.0715	0.2152	0.6216	0.2384	0.2705	0.0023

The weights ($w_{i,k}$) used in equation (a3) are:

j ↓ k →	1
1	10.0470
2	-4.9837
3	4.5256
4	-7.2170
5	4.6952
6	9.1417
7	0.3423