



AUTARQUIA ASSOCIADA À UNIVERSIDADE DE SÃO PAULO

**APLICAÇÃO DE REDES NEURAIS PARA DETECÇÃO DE  
ERROS GROSSEIROS EM DADOS DE PROCESSO DE  
SEPARAÇÃO DE ISÓTOPOS DE URÂNIO POR  
ULTRACENTRIFUGAÇÃO**

**MÔNICA DE CARVALHO VASCONCELOS ANDRADE**

**Dissertação apresentada como parte  
dos requisitos para obtenção do  
Grau de Mestre em Ciências na Área  
de Tecnologia Nuclear - Materiais.**

**Orientadora:  
Dra. Sylvana C.P. Migliavacca**

**São Paulo  
2004**

INSTITUTO DE PESQUISAS ENERGÉTICAS E NUCLEARES

Autarquia associada à Universidade de São Paulo

**APLICAÇÃO DE REDES NEURAIIS PARA DETECÇÃO DE ERROS  
GROSSEIROS EM DADOS DE PROCESSO DE SEPARAÇÃO DE  
ISÓTOPOS DE URÂNIO POR ULTRACENTRIFUGAÇÃO**

MÔNICA DE CARVALHO VASCONCELOS ANDRADE



Dissertação apresentada como parte dos requisitos para obtenção do Grau de Mestre em Ciências na Área de Tecnologia Nuclear-Materiais

Orientador: Dra. Sylvana C. P. Migliavacca

**SÃO PAULO**

**2004**

## AGRADECIMENTOS

Agradeço à Dra. Sylvana Cavedon Presti Migliavacca, minha orientadora, pela correção e orientação deste trabalho, pelo apoio e incentivo, e acima de tudo pela amizade e oportunidade de crescimento profissional que sempre me proporcionou.

Ao Dr. Cláudio Augusto Oller do Nascimento, pelo interesse, pelas valiosas sugestões, e pela permissão de utilização do programa Neuro14.

À FAPESP, pelo auxílio financeiro disponibilizado a este trabalho.

Ao Centro Tecnológico da Marinha em São Paulo (CTMSP), pelo apoio e permissão para o desenvolvimento deste trabalho.

Aos meus colegas do CTMSP-II que, direta ou indiretamente colaboraram para elaboração deste trabalho, de modo especial, a Célia C. Paschoa Portoghese e Elder Migliavacca, pela amizade e incentivo, e pelas discussões sempre úteis relativas ao tema.

Ao meu esposo Delvonei Alves de Andrade pelo apoio pessoal e profissional, incentivo, paciência e principalmente pelo companheirismo e compreensão.

Finalmente agradeço à minha família, meus pais e meus filhos, pelo incentivo e carinho.

## AGRADECIMENTOS

Agradeço à Dra. Sylvana Cavedon Presti Migliavacca, minha orientadora, pela correção e orientação deste trabalho, pelo apoio e incentivo, e acima de tudo pela amizade e oportunidade de crescimento profissional que sempre me proporcionou.

Ao Dr. Cláudio Augusto Oller do Nascimento, pelo interesse, pelas valiosas sugestões, e pela permissão de utilização do programa Neuro14.

À FAPESP, pelo auxílio financeiro disponibilizado a este trabalho.

Ao Centro Tecnológico da Marinha em São Paulo (CTMSP), pelo apoio e permissão para o desenvolvimento deste trabalho.

Aos meus colegas do CTMSP-II que, direta ou indiretamente colaboraram para elaboração deste trabalho, de modo especial, a Célia C. Paschoa Portoghese e Elder Migliavacca, pela amizade e incentivo, e pelas discussões sempre úteis relativas ao tema.

Ao meu esposo Delvonei Alves de Andrade pelo apoio pessoal e profissional, incentivo, paciência e principalmente pelo companheirismo e compreensão.

Finalmente agradeço à minha família, meus pais e meus filhos, pelo incentivo e carinho.

# **APLICAÇÃO DE REDES NEURAIIS PARA DETECÇÃO DE ERROS GROSSEIROS EM DADOS DE PROCESSO DE SEPARAÇÃO DE ISÓTOPOS DE URÂNIO POR ULTRACENTRIFUGAÇÃO**

**Mônica de Carvalho Vasconcelos Andrade**

## **RESUMO**

Neste trabalho é apresentada e discutida a aplicação da técnica de redes neurais para detecção de erros grosseiros em um conjunto de dados experimentais proveniente de ensaios de separação isotópica em ultracentrífugas. A fim de avaliar a aplicação desta nova técnica, o resultado de filtragem obtido por esta foi comparado ao resultado da análise estatística combinado com a análise de "cluster". Este método para detecção de erros grosseiros possui um considerável potencial no campo da análise de dados e é ao mesmo tempo mais fácil e mais rápido para usar e requer muito menos conhecimento da física do processo. Este trabalho estabeleceu um procedimento para indicação de dados suspeitos de conterem erros grosseiros dentro de um conjunto de dados onde as técnicas usuais para identificação destes erros não podem ser aplicadas ou sua utilização demanda um trabalho excessivamente longo.

**DETECTION OF OUTLIERS BY NEURAL NETWORK ON THE GAS  
CENTRIFUGE EXPERIMENTAL DATA OF ISOTOPIC  
SEPARATION PROCESS**

**Mônica de Carvalho Vasconcelos Andrade**

**ABSTRACT**

This work presents and discusses the neural network technique aiming at the detection of outliers on a set of gas centrifuge isotope separation experimental data. In order to evaluate the application of this new technique, the result obtained of the detection is compared to the result of the statistical analysis combined with the cluster analysis. This method for the detection of outliers presents a considerable potential in the field of data analysis and it is at the same time easier and faster to use and requests very less knowledge of the physics involved in the process. This work established a procedure for detecting experiments which are suspect to contain gross errors inside a data set where the usual techniques for identification of these errors cannot be applied or its use demands an excessively long work.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>1</b>
1.1	APRESENTAÇÃO DO TEMA .....	1
1.2	CONCEITOS BÁSICOS DA TEORIA DA SEPARAÇÃO ISOTÓPICA .....	3
1.3	DESCRIÇÃO DO PROCESSO DE CENTRIFUGAÇÃO .....	5
1.4	ERROS EXPERIMENTAIS .....	8
1.5	REDE NEURAL .....	10
<b>2</b>	<b>OBJETIVO .....</b>	<b>14</b>
<b>3</b>	<b>METODOLOGIA.....</b>	<b>15</b>
3.1	OBTENÇÃO DOS DADOS EXPERIMENTAIS .....	15
3.2	BALANÇO MATERIAL MACROSCÓPICO .....	17
3.3	ANÁLISE DE “CLUSTER” .....	17
3.4	ERROS EXPERIMENTAIS .....	18
3.5	ANÁLISE ESTATÍSTICA DOS ERROS EXPERIMENTAIS .....	20
3.6	APLICAÇÃO DE REDE NEURAL EM ULTRACENTRÍFUGAS.....	21
	3.6.1 Primeiro caso: Rede Neural com poder de separação $\delta U$ como variável de saída.....	22
	3.6.2 Segundo caso: Rede Neural com $R_p$ e $R_w$ (razões de abundância) como variáveis de saída .....	23
<b>4</b>	<b>RESULTADOS .....</b>	<b>25</b>
4.1	BALANÇO MATERIAL MACROSCÓPICO .....	25
4.2	ANÁLISE DE “CLUSTER” .....	26
4.3	ANÁLISE ESTATÍSTICA DOS ERROS EXPERIMENTAIS UTILIZANDO TESTE DE HIPÓTESES .....	27
4.4	PRIMEIRO CASO: REDE NEURAL COM PODER DE SEPARAÇÃO $\delta U$ COMO VARIÁVEL DE SAÍDA .....	29
4.5	SEGUNDO CASO: REDE NEURAL COM $R_p$ E $R_w$ (RAZÕES DE ABUNDÂNCIA) COMO VARIÁVEIS DE SAÍDA .....	34
<b>5</b>	<b>DISCUSSÃO DOS RESULTADOS.....</b>	<b>45</b>
5.1	BALANÇO MATERIAL MACROSCÓPICO, ANÁLISE DE CLUSTER E ANÁLISE ESTATÍSTICA .....	45
5.2	PRIMEIRO CASO: REDE NEURAL COM PODER DE SEPARAÇÃO $\delta U$ COMO VARIÁVEL DE SAÍDA .....	46
5.3	SEGUNDO CASO: REDE NEURAL COM $R_p$ E $R_w$ (RAZÕES DE ABUNDÂNCIA) COMO VARIÁVEIS DE SAÍDA .....	47
5.4	VERIFICAÇÃO E VALIDAÇÃO DA METODOLOGIA ATRAVÉS DE UM SEGUNDO CONJUNTO DE DADOS .....	48
<b>6</b>	<b>CONCLUSÕES.....</b>	<b>58</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>61</b>

## LISTA DE FIGURAS

FIGURA 1. 1: Esquema de uma ultracentrífuga a contracorrente (extraída de BENEDICT et al ).....	6
FIGURA 1. 2: Esquema de uma rede neural de três camadas.....	12
FIGURA 3. 1: Esquema de um ensaio de Separação Isotópica na ultracentrífuga a gás.....	16
FIGURA 3. 2: Esquema das variáveis para treinamento da rede neural com $\delta U$ como variável de saída.....	22
FIGURA 3. 3: Esquema proposto para as variáveis de treinamento da rede neural com $R_p$ e $R_w$ como variáveis de saída.....	23
FIGURA 4. 1: Dispersão dos resíduos $Y_{RP}$ com a seqüência dos experimentos.....	27
FIGURA 4. 2: Dispersão dos resíduos $Y_{RW}$ com a seqüência dos experimentos.....	28
FIGURA 4. 3: Erro global das redes com 6, 8 e 10 neurônios com treinamento total até 50000 apresentações - variável de saída: $\delta U$ .....	30
FIGURA 4. 4: Erro global das redes com 6, 8 e 10 neurônios com treinamento total até 5000 apresentações - variável de saída: $\delta U$ .....	31
FIGURA 4. 5: Comparação entre os valores experimentais e calculados do poder de separação $\delta U$ pela rede neural de 6 neurônios na camada oculta com 5000 apresentações.....	32
FIGURA 4. 6: Comparação entre os valores experimentais e calculados do poder de separação $\delta U$ pela rede neural de 8 neurônios na camada oculta com 5000 apresentações.....	32
FIGURA 4. 7: Erro global das redes com 6, 8 e 10 neurônios com treinamento total até 5000 apresentações - variáveis de saída: $R_p$ e $R_w$ .....	35
FIGURA 4. 8: Análise dos resíduos para a variável de saída $R_p$ .....	36
FIGURA 4. 9: Análise dos resíduos para a variável de saída $R_w$ .....	36
FIGURA 4. 10: Erro global no treinamento da rede neural com 8 neurônios na camada oculta para o primeiro subconjunto.....	37
FIGURA 4. 11: Erro global no treinamento da rede neural com 8 neurônios na camada oculta para o segundo subconjunto.....	38
FIGURA 4. 12: Erro global no treinamento da rede neural com 8 neurônios na camada oculta para o terceiro subconjunto.....	38
FIGURA 4. 13: Análise dos resíduos para a variável de saída $R_p$ para o 1º subconjunto.....	39
FIGURA 4. 14: Análise dos resíduos para a variável de saída $R_w$ para o 1º subconjunto.....	39
FIGURA 4. 15: Análise dos resíduos para a variável de saída $R_p$ para o 2º subconjunto.....	40
FIGURA 4. 16: Análise dos resíduos para a variável de saída $R_w$ para o 2º subconjunto.....	40
FIGURA 4. 17: Análise dos resíduos para a variável de saída $R_p$ para o 3º subconjunto.....	41
FIGURA 4. 18: Análise dos resíduos para a variável de saída $R_w$ para o 3º subconjunto.....	41
FIGURA 5. 1: Esquema das variáveis utilizadas na apresentação à rede neural.....	50
FIGURA 5. 2: Erro total das redes com 6, 8 e 10 neurônios ao longo de seus treinamentos até 50000 apresentações.....	51
FIGURA 5. 3: Erro total das redes com 6, 8 e 10 neurônios ao longo de seus treinamentos até 5000 apresentações.....	52



## LISTA DE TABELAS

TABELA 4. 1 : Relação dos experimentos rejeitados por insuficiência de material para análise por espectrometria de massa .....	25
TABELA 4. 2 : Relação dos experimentos que não respeitam o balanço material macroscópico.....	26
TABELA 4. 3 : Experimentos considerados suspeitos pela análise de cluster.....	27
TABELA 4. 4 : Teste de Hipóteses dos erros experimentais.....	28
TABELA 4. 5 : Relação dos experimentos considerados suspeitos pela análise estatística.....	29
TABELA 4. 6 : Erro global mínimo encontrado para cada rede neural aplicada permitidas até 50000 apresentações no treinamento .....	30
TABELA 4. 7 : Teste de Hipóteses dos resultados das redes neurais com 6 neurônios na camada oculta ....	33
TABELA 4. 8 : Teste de Hipóteses dos resultados da rede neural com 8 neurônios na camada oculta para 5000 apresentações.....	34
TABELA 4. 9 : Teste de Hipóteses dos resultados das redes neurais com 8 neurônios na camada oculta para o 1o subconjunto de dados .....	42
TABELA 4. 10 : Teste de Hipóteses dos resultados das redes neurais com 8 neurônios na camada oculta para o 2o subconjunto de dados.....	42
TABELA 4. 11 : Teste de Hipóteses dos resultados das redes neurais com 8 neurônios na camada oculta para o 3o subconjunto de dados.....	43
TABELA 4. 12 : Indicação de dados suspeitos de conterem erros grosseiros pelas redes neurais para os 3 subconjuntos de dados após cada rodada .....	44
TABELA 5. 1 : Distribuição e Identificação dos ensaios utilizados no segundo conjunto de dados.....	49
TABELA 5. 2 : Erro total mínimo encontrado para cada rede neural aplicada permitidas até 50000 apresentações no treinamento .....	51
TABELA 5. 3 : Identificação de possíveis erros grosseiros através de diferentes redes neurais sendo o número de neurônios na camada oculta NH = 6 Primeira rodada .....	53
TABELA 5. 4 : Identificação de possíveis erros grosseiros através de diferentes redes neurais sendo o número de neurônios na camada oculta NH = 8 Primeira rodada .....	53
TABELA 5. 5 : Identificação de possíveis erros grosseiros através de diferentes redes neurais sendo o número de neurônios na camada oculta NH = 10 Primeira rodada .....	54
TABELA 5. 6 : Identificação de possíveis erros grosseiros através de diferentes redes neurais sendo o número de neurônios na camada oculta NH = 6 Segunda rodada.....	54
TABELA 5. 7 : Identificação de possíveis erros grosseiros através de diferentes redes neurais sendo o número de neurônios na camada oculta NH = 8 Segunda rodada.....	55
TABELA 5. 8 : Identificação de possíveis erros grosseiros através de diferentes redes neurais sendo o número de neurônios na camada oculta NH = 10 Segunda rodada.....	55
TABELA 5. 9 : Identificação de possíveis erros grosseiros através de diferentes redes neurais.....	56

# 1 INTRODUÇÃO

## 1.1 Apresentação do tema

Qualquer conjunto de dados experimentais apresenta um erro associado, seja de origem sistemática, randômica ou, na maior parte dos casos, ambas. Pode-se definir um erro como a discrepância entre um valor observado e o valor verdadeiro. Porém, dificilmente temos o valor verdadeiro, mas somente uma estimativa que seria a média de um número finito de observações. Os erros das medidas resultam de uma série de fatores não perfeitamente controlados e que atuam sobre os resultados de modo praticamente independente uns dos outros. Algumas mudanças nas condições do processo durante um experimento podem permanecer não detectadas, ou mesmo medidas estranhas podem ocorrer devido a erros na operação de instrumentos, dispositivos de registro, gerando assim, os chamados erros grosseiros. A presença de medidas indesejáveis no conjunto de dados deteriora a estimativa efetuada, conseqüentemente, a modelagem do processo. Porém, tais erros precisam ser identificados e investigados segundo alguns critérios de detecção, antes de serem excluídos do conjunto de dados, de forma a não permitir que observações corretas sejam rejeitadas.

O uso adequado desses dados experimentais necessita uma cuidadosa análise dos erros associados, que indicam a reprodutibilidade, a representatividade e a confiabilidade dos dados obtidos. No entanto, esse tipo de análise pode ser uma tarefa árdua quando se trata de um processo muito complexo, onde parâmetros não controlados ou não monitorados influenciam o processo, ou onde se observa grandes desvios estocásticos, resultando em um conjunto de dados muito disperso.

Existem vários procedimentos propostos para o tratamento e análise de dados experimentais, baseados em estatística ou na física do processo (Himmelblau,1970 e Plácido,1995). No entanto essas técnicas podem ser inviáveis em alguns casos, por exemplo quando a física do processo não é bem conhecida (isto é, não pode ser totalmente representada por meio de equações)

ou quando critérios estatísticos tradicionais não podem ser aplicados para a separação dos dados de um conjunto em valores válidos e não válidos (por exemplo, quando existem poucos dados em uma mesma condição de operação).

Muitos processos apresentam essas características, como o processo de investigação e desenvolvimento do enriquecimento de urânio por ultracentrifugação (Migliavacca,1999b), onde a estimativa numérica do desempenho separativo de uma ultracentrífuga passa pelo cálculo do escoamento interno envolvendo o estudo de gás-dinâmica, cuja solução só é viável através de simplificações, o que freqüentemente leva a soluções distantes da realidade.

A aplicação de redes neurais na análise de processos altamente não lineares tem se mostrado uma poderosa ferramenta. Este trabalho propõe a utilização de redes neurais para a filtragem de dados experimentais provenientes desses processos complexos. A viabilidade deste tipo de aplicação e seu grande potencial são apresentados por Bülow et al. (1999) apontando a maior facilidade e agilidade em relação às técnicas tradicionais de reconciliação de dados, no entanto este trabalho não estabelece critérios mais gerais.

## 1.2 Conceitos Básicos da Teoria da Separação Isotópica

A separação dos isótopos do urânio, assim como a separação de outros isótopos estáveis, independentemente do processo utilizado, pode ser representada por uma teoria completamente geral que se aplica a qualquer mistura isotópica binária. Uma vez que as propriedades físicas e químicas dos isótopos são muito próximas, a separação realizada em uma única vez é, de modo geral, muito pequena. Por isso se faz necessária a ligação em série de vários dispositivos de separação com o objetivo de multiplicar a separação obtida. A esse conjunto chamamos de cascata de separação de isótopos. A teoria da separação dos isótopos do urânio assim realizada está descrita em detalhes nas obras de Cohen (1951) e Benedict et al. (1981) e na obra nacional de Jordan (1980).

O menor elemento de uma usina de enriquecimento é chamado unidade ou elemento de separação. No caso da separação do urânio pela ultracentrífuga a gás, uma unidade de separação é representada por uma ultracentrífuga. Um grupo de unidades de separação ligadas em paralelo define um estágio. No estágio todas as unidades de separação são alimentadas com material de mesma composição isotópica e produzem individualmente frações enriquecidas, ou produtos, e frações empobrecidas, ou rejeitos, com composições iguais entre si. A ligação em série de um grupo de estágios forma então uma cascata de separação isotópica.

O elemento de separação simples recebe um fluxo de alimentação  $F$  da mistura isotópica binária de composição  $x_F$  do isótopo desejado ( $^{235}\text{U}$ ) e fornece duas correntes, uma, a do produto, parcialmente enriquecida no isótopo desejado com fluxo  $P$  e composição  $x_P$  e outra parcialmente empobrecida no isótopo desejado, denominada rejeito, de fluxo  $W$  e composição  $x_W$ . A razão de abundância, definida pela relação entre as frações atômicas do componente desejado e a do outro componente  $R = x/(1-x)$  é denotada por  $R_F$  para a corrente da alimentação,  $R_P$  para a corrente do produto e  $R_W$  para a do rejeito. Definimos a seguir os principais parâmetros de separação:

$$\text{fator de enriquecimento: } \beta = \frac{R_P}{R_F} = \frac{x_P(1-x_F)}{x_F(1-x_P)} \quad (1.1)$$

$$\text{fator de empobrecimento: } \gamma = \frac{R_F}{R_W} = \frac{x_F(1-x_W)}{x_W(1-x_F)} \quad (1.2)$$

$$\text{fator de separação: } \alpha = \frac{R_P}{R_W} = \frac{x_P(1-x_W)}{x_W(1-x_P)} = \beta \times \gamma \quad (1.3)$$

Chamamos de processo simétrico o processo no qual os fatores de enriquecimento e de empobrecimento são iguais,  $\beta=\gamma$ , e portanto,  $\alpha=\beta^2$ , e denotamos os parâmetros para este caso por um asterisco (\*).

Outra propriedade fundamental da unidade de separação é o corte  $\theta$ . Essa propriedade é definida como sendo a relação entre o fluxo da corrente enriquecida e o fluxo da alimentação:

$$\theta = \frac{P}{F} = \frac{P}{P+W} \quad (1.4)$$

ou, em termos das composições, quando se considera o balanço material do isótopo desejado:

$$\theta = \frac{x_F - x_W}{x_P - x_W} \quad (1.5)$$

A unidade de separação é descrita pelos parâmetros separativos  $\alpha$  ou  $\beta$  e pelas variáveis de fluxo  $F$  e  $\theta$  (ou  $P$  e  $W$ ). No entanto, nenhuma dessas variáveis representa individualmente o desempenho da unidade de separação. Assim, uma propriedade que englobe as grandezas separativas e de fluxo e que ao mesmo tempo esteja relacionada ao trabalho separativo útil produzido pela unidade, torna-se atrativa. Essa propriedade, chamada poder ou capacidade de separação, foi definida por Dirac, apresentada num manuscrito em 1941 não publicado, mas citado por Cohen (1951), mediante a introdução do conceito do fluxo de valor, considerando que a unidade de separação produz uma variação no valor de uma quantidade de material separado. Sendo o fluxo de valor uma função  $U$ , a produção de valor resultante da separação de isótopos por um elemento de separação é dada pela variação da função  $U$ . O fluxo de valor  $U$  é dado pelo produto do fluxo de material e uma função de valor  $V(x)$ . Então no elemento de separação vamos ter os fluxos de valor  $U_F=FV(x_F)$ ,  $U_P=PV(x_P)$  e

$U_w = WV(x_w)$  para as correntes de alimentação, produto e rejeito, respectivamente. A variação do valor produzida pelo elemento de separação é então o poder de separação  $\delta U$ , o qual é dado pelo balanço dos fluxos de valor, a saber:

$$\delta U = PV(x_p) + WV(x_w) - FV(x_f) \quad (1.6)$$

A função de valor  $V(x)$  é uma função somente da composição  $x$  e é adimensional, definida pela equação:

$$V(x) = (2x - 1) \ln\left(\frac{x}{1-x}\right) \quad (1.7)$$

Esta função é simétrica em  $x=0,5$ , onde ela assume valor zero. É positiva para os demais valores de composição e aumenta tendendo ao infinito quando  $x$  se aproxima de 0 e de 1. A função  $V(x)$  permite que o aumento de valor  $\delta U$  efetuado pela unidade de separação seja independente da concentração isotópica do material processado. O poder de separação  $\delta U$  foi adotado internacionalmente como parâmetro para quantificar a capacidade de unidades e instalações e é empregado em trabalhos de separação de isótopos de urânio. A explicação do significado físico desta função não é trivial. Guangjun (1994) mostrou que a função valor pode ser interpretada como uma quantidade proporcional ao grau de ordem do sistema, isto feito através de análise da entropia do sistema.

### 1.3 Descrição do Processo de Centrifugação

A ultracentrífuga esquematizada na Figura 1.1, do tipo de fluxo em contracorrente, consiste de um cilindro vertical relativamente longo e de parede fina (rotor), girando em altas velocidades dentro de um recipiente (carcaça) mantido em alto vácuo. O gás de processo, uma mistura isotópica binária contendo  $^{235}\text{UF}_6$  e  $^{238}\text{UF}_6$ , no interior do rotor está sujeito à ação de uma aceleração centrífuga milhares de vezes maior que a aceleração da gravidade. Assim se estabelece uma distribuição de pressões, manifestada na forma de um gradiente radial de pressão, crescente no sentido do eixo para a parede do rotor.

Essa distribuição de pressão é diferente para os diferentes isótopos, uma vez que a ação do campo centrífugo depende da massa. Resulta então uma separação parcial da mistura de isótopos presente no interior do rotor. Induzindo um fluxo em contracorrente no sentido vertical, multiplica-se o efeito da separação radial, o que faz com que a diferença entre a composição do topo e do fundo do rotor da ultracentrífuga se torne maior que a diferença no sentido radial para uma dada posição axial.

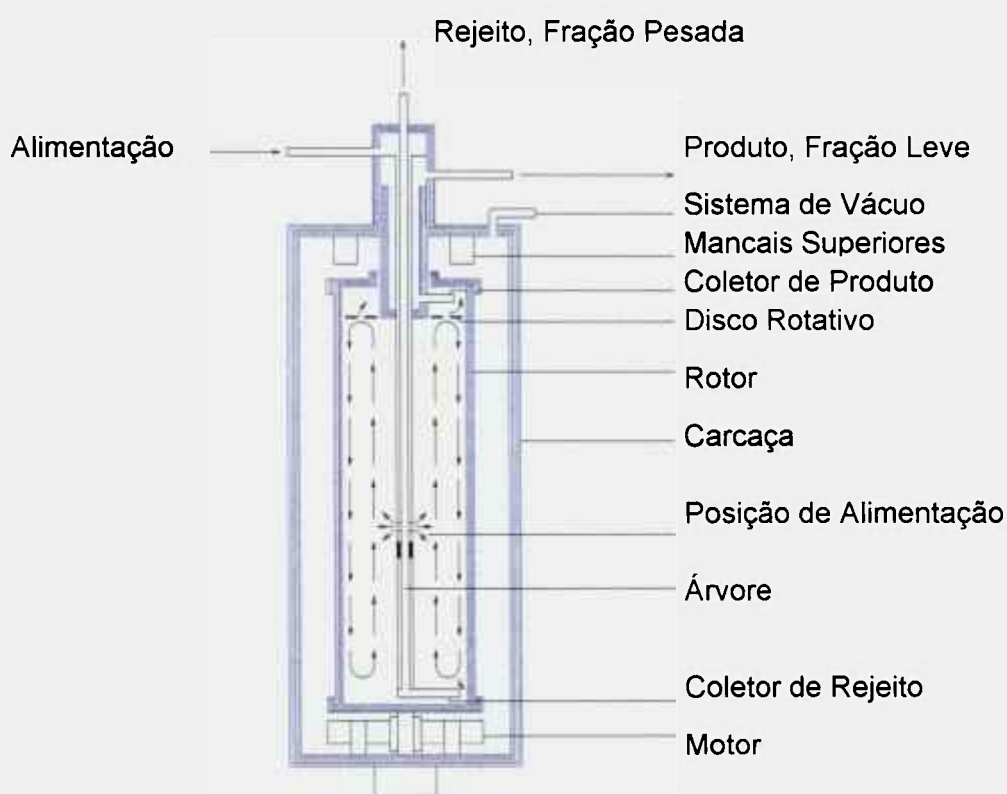


FIGURA 1. 1: Esquema de uma centrífuga a contracorrente (extraída de Benedict et al )

O desempenho separativo de uma ultracentrífuga (Avery e Davis, 1973; Benedict et al., 1981) depende: (a) do tipo da ultracentrífuga, definido pela velocidade de rotação, comprimento e diâmetro do rotor; (b) das condições de operação dadas pelos fluxos de alimentação e retiradas do gás de processo e pela quantidade de material no interior da ultracentrífuga (ou pela pressão de operação); (c) pela posição axial na qual a alimentação é introduzida no rotor e; (d) pelo perfil e intensidade do fluxo em contracorrente. Podem existir três fatores

responsáveis pela geração da circulação interna em contracorrente, representada pelas setas verticais na Figura 1.1: (1) a injeção do fluxo de alimentação no interior do rotor; (2) a presença de um coletor ou um outro obstáculo estacionário em uma das extremidades do rotor e uma placa rotativa na outra extremidade; (3) por correntes de convecção produzidas por uma diferença entre as temperaturas das tampas do rotor. Na prática, esses três fatores são responsáveis pela geração da contracorrente. As equações nas quais se baseia a teoria da ultracentrífuga a contracorrente, para a separação isotópica do urânio, são as seguintes:

#### **Equações de conservação do gás como um todo**

1. Massa: equação da continuidade
2. Quantidade de movimento radial
3. Quantidade de movimento angular
4. Quantidade de movimento axial
5. Energia

#### **Equação termodinâmica de estado**

6. Equação do gás ideal

#### **Equação de conservação da mistura gasosa**

7. Massa: equação da continuidade do isótopo desejado

Nesse caso, os perfis de circulação interna, resultantes de cada um desses fatores, podem ser analisados separadamente quando são admitidas hipóteses simplificadoras na solução das seis primeiras equações. Por outro lado, a obtenção dos parâmetros separativos da ultracentrífuga, através do cálculo da distribuição de concentração dos isótopos, não permite a linearização da sétima equação que descreve o equilíbrio entre os dois componentes ( $^{235}\text{UF}_6$  e  $^{238}\text{UF}_6$ ), também conhecida como equação da difusão-convecção. Para essa análise é necessário recorrer à solução das equações que descrevem o fluxo total no interior do rotor.

Assim, a estimativa numérica do desempenho separativo de uma ultracentrífuga passa pelo cálculo do escoamento interno, cuja solução só é viável através de simplificações (Kai, 1989; Migliavacca, 1991; Olander, 1972 e



1981; Soubbaramayer, 1979). Por outro lado, é muito difícil relacionar os parâmetros utilizados para descrever numericamente a contracorrente com os dispositivos físicos empregados na ultracentrífuga (Kai, 1989; Migliavacca, 1999a e 1999b). Por se tratar de um processo onde o gás escoar em velocidades supersônicas, o escoamento resultante é muito sensível a variações das características físicas da ultracentrífuga, chegando até mesmo a sentir diferenças dentro da precisão mecânica definida.

#### **1.4 Erros experimentais**

Experimentos planejados e executados cuidadosamente podem resultar em dados não homogêneos. Mudanças nas condições durante um experimento podem permanecer não detectadas, tal que freqüentemente ocorram medidas anormais, os chamados valores grosseiros (Himmelblau, 1970).

Os erros randômicos são comuns a todo processo experimental de obtenção de dados, constituindo pequenos desvios em relação a um valor central (são gaussianos).

Os erros ditos grosseiros constituem valores que ultrapassam a faixa dos erros gaussianos, e são conseqüência de diversos fatores.

Erros na operação de dispositivos de registro podem causar algumas medidas aberrantes, o que resultaria na rejeição dos valores registrados, caso essas medidas fossem conhecidas.

Mesmo dados válidos podem ser contaminados por dígitos invertidos ou pontos decimais deslocados, que são os chamados erros de digitação.

Na prática é muito difícil que exista um processo verdadeiramente em regime permanente. Existem perturbações naturais ou provocadas, de origens diversas e, também, insuspeitas, tais como as oriundas das ações de controle, mesmo que mínimas; alterações momentâneas nos percentuais da carga do processo; flutuações de pressões inesperadas; revezamento entre bombas podendo provocar flutuações de vazão; ações corriqueiras dos operadores, por vezes inadequadas, causando, por exemplo, bloqueio momentâneo de correntes.

Deste modo é ilusório concluir, por exemplo, que as medições levantadas são boas com base apenas no fato de que os medidores de vazão, ou de pressão, tem tal ou qual precisão (Plácido, 1995). As condições ambientais também podem ter influência significativa nas leituras em equipamentos sob temperaturas que não estão de conformidade com as especificações dos fabricantes.

Em resumo, dentre as causas de erro grosseiro, pode-se citar:

- descalibração ou calibração deficiente dos instrumentos de medição (que provocam desvios – “bias”);
- vazamento insuspeito nas unidades do processo;
- mal funcionamento ou falha dos instrumentos;
- regime transiente ou flutuações naturais do processo;
- flutuações nas condições ambientais;
- erros de amostragem e de análise;
- leituras feitas por diferentes observadores;
- dados levantados em tempos não simultâneos, pertencentes a regimes permanentes distintos, etc.

Por outro lado, os erros identificados como grosseiros podem ser simplesmente um dos valores extremos em uma distribuição de probabilidade para uma variável randômica a qual ocorre muito naturalmente, mas não freqüentemente e deveria não ser rejeitada.

Os tratamentos estatísticos mais comuns assumem que as medições contêm apenas erros randômicos, em geral considerados independentes e normalmente distribuídos com média zero (Plácido, 1995). Porém, o conjunto de dados medidos pode conter erros grosseiros, os quais deterioram a estimativa efetuada, sendo necessário um algoritmo específico para a detecção e remoção destes erros.

Quando o analista sabe que um erro anormal foi feito, ele não hesita em descartar tal observação. Quando ele não tem bases práticas suficientes para também aceitar ou rejeitar uma observação extrema, deve recorrer a algum tipo de julgamento estatístico. Ele gostaria de responder a questão: “Qual a probabilidade de que as diferenças observadas sejam devido somente a erros amostrais randômicos?”, de tal modo que haja pequena dúvida que observações corretas sejam rejeitadas (Himmelblau, 1970).

A abordagem do problema de análise de observações contendo erros grosseiros depende acima de tudo do objetivo em mãos. Se o analista está interessado somente em determinar se uma observação é candidata a erro grosseiro, talvez, para investigar a condição ou condições que possam ter conduzido a esta observação extrema, então o teste para tal observação de erro grosseiro é um fim nele mesmo. Se por outro lado, ele está interessado em excluir os candidatos a erros grosseiros para obter uma estimativa mais precisa de alguns parâmetros da população, ou seja, média da população, então ele não está interessado somente em um teste de determinação de erros grosseiros, mas também na estimativa de parâmetros subsequentes do teste.

Usualmente, a combinação do erro instrumental, as equações físicas básicas do processo considerado, e algum tipo de tratamento estatístico é usado para detectar e filtrar estes erros grosseiros. Entretanto, se o controle ou mesmo a monitoração de algumas variáveis influentes do processo é impossível, dificuldades consideráveis podem surgir quando tal tratamento de dados é aplicado. Por esta razão, um novo caminho é proposto com o uso de redes neurais para detecção de erros grosseiros (Bülau et al, 1999).

## 1.5 Rede Neural

A técnica de redes neurais é uma das áreas de inteligência artificial que mais se desenvolveu nos últimos anos. Na engenharia química só recentemente essa técnica começou a ser utilizada, notadamente em detecção e diagnóstico de falhas de processos químicos (Hoskins et al., 1988 e 1990, Venkatasubramanian et al. 1990, Watanabe et al. 1989) e em modelagem e controle de processos químicos (Bhat et al. 1990, Bhat & McAvoy, 1990, Ungar et al., 1990, Psychogios & Ungar, 1991, Su & McAvoy, 1992, Chan & Nascimento, 1994).

Na tecnologia nuclear essa técnica começou a ser explorada em 1989 (Yiftah, 1989) e tem sido largamente utilizada em física de alta energia (Denby, 1992, Cutts, 1990 e Peterson, 1990) e em usinas nucleares (Uhrig, 1991a, 1991b,

1991c, Tanabe, 1991, Eryürek et.al., 1991a e 1991b). Nesse último campo vale a pena citar as possíveis aplicações de redes neurais:

- (a) diagnóstico de condições anormais específicas;
- (b) detecção de mudanças no modo de operação;
- (c) validação de sinal de instrumentos;
- (d) monitoração da checagem de válvulas;
- (e) monitoração total da planta usando redes neurais auto-associativas;
- (f) modelagem termodinâmica da planta;
- (g) emulação dos cálculos da recarga do núcleo;
- (h) análise de vibrações da planta;
- (i) estimativa do nível de potência durante um transiente de desligamento da planta e em condições normais de operação;
- (j) monitoração do nível de potência e das temperaturas da “perna quente”, da “perna fria” e da saída do núcleo do reator.

Existem duas estruturas principais de redes neurais: (i) rede de múltiplas camadas, cuja informação segue uma única direção ("*multilayer feedforward network*"), esquematizada na Figura 1.2, utilizadas para modelagem de sistemas em estado estacionário, e (ii) rede recursiva ("*recurrent network*"), para previsão de parâmetros, e portanto mais apropriada para modelos dinâmicos.

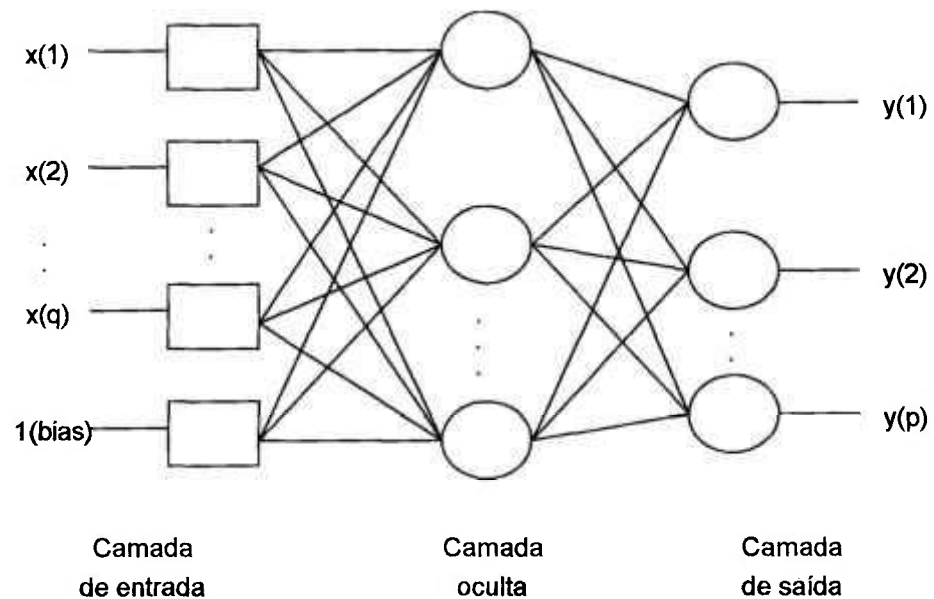


FIGURA 1. 2: Esquema de uma rede neural de três camadas

Em geral, uma rede é formada por neurônios (representados por círculos), e por canais que transferem a informação entre os neurônios, denominados interconectores. Os retângulos representam neurônios onde as entradas da rede são armazenadas. Cada neurônio calcula primeiramente a soma ponderada de todos os sinais da camada anterior (Eq.1.8), onde  $W_{u,v}$  são os pesos entre a camada de entrada e a camada oculta, e então gera uma saída através de uma função, geralmente definida pela Eq.1.9.

$$S_v = \sum_{u=1}^n W_{u,v} x_u + W_{n+1,v} \quad (1.8)$$

$$f(S_v) = \frac{1}{1 + e^{-S_v}} \quad (1.9)$$

As redes neurais mais comumente utilizadas para modelagem de processos são compostas por três camadas: a camada de entrada de dados, a camada oculta ("*hidden layer*") e a camada de saída. A camada de entrada possui  $n_i+1$  neurônios, onde  $n_i$  é o número de variáveis de entrada, e não efetua nenhum processamento. A fim de diminuir as dificuldades de convergência, todos os dados do conjunto de aprendizagem são normalizados entre zero e um. Além dos dados de entrada, a rede também é alimentada com um valor de referência

("bias"). O número de neurônios da camada oculta é definido pelo usuário, e, segundo Pollard et al., 1992, a precisão dos resultados finais é relativamente insensível à quantidade de neurônios, acima de um valor mínimo necessário. A camada de saída é constituída por um número de neurônios igual ao número de variáveis de saída ou respostas.

O sistema "aprende" alterando os valores dos pesos ( $W_{u,v}$ ) de forma a minimizar o quadrado das diferenças ( $E$ ) entre os valores experimentais e os calculados pela rede, definido por

$$E = \sum_{m=1}^1 \sum_{k=1}^{no} (y_k^{(m)} - O_k^{(m)})^2 \quad (1.10)$$

onde  $y_k$  é proveniente do conjunto de dados experimentais para o treinamento da rede e  $O_k$  é o valor correspondente obtido na camada de saída calculado pela expressão

$$O_k = f(S_k) \quad (1.11)$$

Atualmente, o algoritmo mais utilizado para determinar o conjunto de pesos no aprendizado da rede é o de retro-propagação ("backpropagation"). Esse algoritmo é uma derivação do gradiente descendente (Rumelhart & McClelland, 1986).

## 2 OBJETIVO

O objetivo do presente trabalho é a detecção de erros grosseiros através da aplicação da técnica de redes neurais em um conjunto de dados experimentais proveniente de ensaios de separação isotópica em ultracentrífugas. A fim de avaliar a aplicação desta nova técnica, o resultado de filtragem obtido por ela será comparado ao resultado da análise estatística combinado com a análise de "cluster". Esses dados são originados do processo de avaliação e otimização das ultracentrífugas desenvolvidas pelo Centro Tecnológico da Marinha em São Paulo (CTMSP), com a colaboração do Instituto de Pesquisas Energéticas e Nucleares (IPEN).

Como será mostrado, este método para detecção de erros grosseiros possui um considerável potencial no campo da análise de dados e é ao mesmo tempo mais fácil e mais rápido para usar e requer muito menos conhecimento da física do processo.

## 3 METODOLOGIA

### 3.1 Obtenção dos dados experimentais

No presente trabalho foram utilizados dados resultantes de ensaios de separação isotópica por ultracentrifugação. Esses dados são originários do processo de avaliação e otimização das ultracentrífugas desenvolvidas pelo Centro Tecnológico da Marinha em São Paulo (CTMSP). Em um ensaio de separação isotópica de uma ultracentrífuga se estabelece um fluxo do gás de processo  $UF_6$ , de composição isotópica conhecida, através da ultracentrífuga. A operação é efetuada em diversas condições de vazão e pressão nas linhas de alimentação e retiradas de produto e rejeito (material enriquecido em  $^{235}U$  e material empobrecido em  $^{235}U$ , respectivamente). Resultados experimentais semelhantes, guardadas as diferenças decorrentes do avanço tecnológico, foram publicados por Zippe (1960) e Jordan (1966).

Os ensaios são realizados em Estações de Ensaio Instrumentadas (EEI) que são constituídas por tubulações interligadas por válvulas, onde instrumentos e válvulas de controle são apropriadamente localizados de modo a controlar e monitorar todo o processo de introdução e retirada do gás de processo  $UF_6$  na ultracentrífuga em estudo. Uma condição de operação é definida pelas pressões de operação nas linhas de alimentação  $p_F$ , de retirada do produto  $p_P$  e de retirada do rejeito  $p_W$ , pelos fluxos de massa da alimentação  $F$ , de extração do produto  $P$  e do rejeito  $W$ , ou pela relação de corte  $\theta$ . A cada condição de operação estabelecida, são retiradas amostras para verificação da separação obtida através das medidas das razões de abundância das correntes enriquecida e empobrecida,  $R_P$  e  $R_W$  respectivamente, como é mostrado no esquema da Figura 3.1.

Um ensaio de separação isotópica consiste na operação de uma ultracentrífuga em diferentes condições de operação, definidas por três das variáveis do processo (geralmente o fluxo de alimentação  $F$ , a pressão de



retirada do produto  $p_p$  e o corte  $\theta$ ). Logo, são gerados diversos conjuntos de dados, denominados de experimentos de separação, compostos pelas variáveis de controle que definem a condição de operação, pelas demais variáveis do processo, que são então as respostas observadas e os respectivos parâmetros de separação calculados através das Eqs.(1.1) a (1.6). Esses experimentos podem ser agrupados em experimentos obtidos em condições semelhantes, denominados por séries de experimentos.

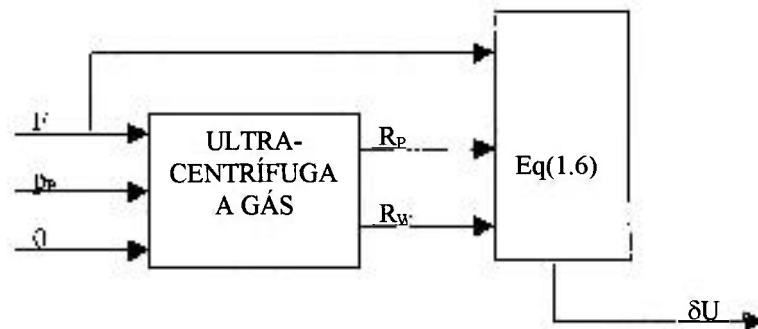


FIGURA 3. 1: Esquema de um ensaio de Separação Isotópica na ultracentrífuga a gás

O primeiro conjunto de dados a ser tratado consta de um ensaio realizado com uma ultracentrífuga, que gerou 791 experimentos. Nesse ensaio foram tomadas como variáveis de controle o fluxo de alimentação  $F$ , o corte  $\theta$  e a pressão na retirada do produto  $p_p$ . O ensaio foi programado de modo a cobrir todo o domínio de interesse dessas variáveis, constando de 8 níveis de fluxo de alimentação  $F$ , 7 níveis de corte  $\theta$  e 5 níveis de pressão  $p_p$ . Dos 791 experimentos realizados, 27 apresentaram condições insuficientes para realização de análise por espectrometria de massa da amostra. Portanto foi necessário desconsiderá-los antes mesmo de uma análise detalhada de dados, resultando em um conjunto de 764 experimentos. Porém, os números dos experimentos foram preservados na ordem de execução destes dentro do ensaio.

Esses dados foram organizados e analisados segundo procedimentos estatísticos convencionais, levando em consideração o balanço de massa macroscópico, a análise de “clusters” e a análise estatística dos erros dos

resultados experimentais, que são discutidos a seguir. Posteriormente, foi empregada a técnica de redes neurais, com diferentes configurações das variáveis e do número de neurônios.

Um segundo conjunto de dados foi utilizado para verificação e validação da metodologia desenvolvida. Este conjunto é formado por dados adquiridos em dez diferentes ensaios realizados com diferentes ultracentrífugas, com o objetivo de avaliar a influência de características de montagem no desempenho separativo das mesmas.

### 3.2 Balanço Material Macroscópico

Durante o acompanhamento do ensaio de separação isotópica, rotineiramente é realizada uma verificação do balanço de massa macroscópico, na comparação do corte obtido através das medidas dos fluxos de produto e de rejeito  $\theta = P/(P+W)$ , proveniente do balanço de massa da mistura, e do corte obtido a partir das razões de abundância  $\theta = \frac{(R_F - R_W)}{(R_P - R_W)} \cdot \frac{(R_P + 1)}{(R_F + 1)}$ , resultante do

balanço de massa do isótopo leve. Experimentos onde a diferença entre esses dois valores de corte é maior que 0,03 são considerados experimentos suspeitos e devem ser repetidos ainda durante o ensaio. Os desvios nas duas determinações do corte  $\theta$  são decorrentes da possibilidade de introdução de erro no procedimento de coleta de amostras através de ampolas, que além da utilização de um elemento externo (a ampola), procede um desvio do fluxo de material que sai da ultracentrífuga para o sistema de amostragem, o que freqüentemente pode perturbar a condição de estado estacionário da ultracentrífuga.

### 3.3 Análise de “Cluster”

Análise de “clusters” é uma técnica de análise multivariável que tem como objetivo agrupar informações baseando-se nas características que possuem. A análise via “clusters” classifica dados de maneira que esses apresentem similaridades com respeito a critérios pré-determinados. Os “clusters”

devem apresentar uma grande homogeneidade entre seus elementos e apresentar uma grande heterogeneidade entre "clusters". Esta análise é a única técnica multivariável cuja importância das variáveis não é estimada empiricamente mas especificada pelo pesquisador/engenheiro. A escolha das variáveis é um passo crítico na análise de "clusters".

É uma técnica bastante útil quando se deseja desenvolver hipóteses com respeito a natureza dos dados. Embora a análise via "clusters" apresente grandes benefícios, ela não possui uma base estatística na qual se possa obter inferências estatísticas. Ela é primariamente uma técnica exploratória. A retirada ou a adição de variáveis importantes de um "cluster" pode ter um grande impacto nos resultados, o que torna cada decisão uma tarefa que deve ser analisada cuidadosamente.

Hair et al (1998) cita que a análise via "clusters" é mais arte do que ciência, isto significa que a sistematização dessa técnica não é trivial uma vez que as principais decisões de escolha da importância relativa de cada variável deve ser do pesquisador/engenheiro.

A análise de "clusters" é sensível à escolha de variáveis irrelevantes, assim como à existência de erros grosseiros nos dados.

Em ambos os casos, os conjuntos de dados foram submetidos a uma análise por "clusters". Para tal, cada um deles foi subdividido em grupos onde duas das variáveis de controle eram mantidas constantes, e uma resposta, no caso o poder de separação  $\delta U$  por ser a mais significativa, era analisada graficamente em função da terceira variável de controle. No segundo conjunto de dados, adicionalmente foram avaliados grupos com características similares de montagem, ou seja, os ensaios foram separados dois a dois, conforme suas similaridades.

### **3.4 Erros Experimentais**

Através do balanço macroscópico no conjunto de dados original do primeiro caso, foi feita uma pré-filtragem destes dados, retirando-se os que não se adequavam ao critério preestabelecido. Entretanto, a necessidade de uma análise estatística ainda existia.

Os tratamentos estatísticos mais comuns assumem que as medições contêm apenas erros randômicos, em geral considerados independentes e normalmente distribuídos com média zero. A presença de erros grosseiros no conjunto de dados medidos deteriora a estimativa efetuada.

Como não fosse possível tratar todos os dados como sendo pertencentes a uma mesma população, da forma como eram apresentados, por serem obtidos em condições de operação diferentes, decidiu-se fazer uma mudança de variáveis, a fim de que as variações das condições de processo não interferissem. Dessa forma, considerou-se a possibilidade de se usar as médias das observações de condições similares de processo. Os erros das medidas resultam de uma série de fatores não perfeitamente controlados e que atuam sobre os resultados. Logo, resolveu-se trabalhar com as variáveis primárias  $R_P$  e  $R_W$  em lugar do  $\delta U$ , tentando-se minimizar a influência de certos fatores indesejáveis durante o processo.

Portanto, a análise estatística dos dados experimentais originais foi baseada na “média das séries”: resultados de mesma condição de processo, fornecendo um total de 158 séries no primeiro conjunto de dados analisado. As médias foram calculadas com as variáveis primárias de processo  $R_P$  e  $R_W$ . Depois, calculou-se o erro do dado experimental em relação à média da série a que ele pertença. Por exemplo, se a série 1 possui 3 amostragens (3 experimentos), tem-se uma média para  $R_P$  igual a  $R_{Pm1}$ , e  $e_{Rp1}$  é a diferença entre  $R_P$  e  $R_{Pm1}$ , tal que:

$$R_{P1} - R_{Pm1} = e_{Rp1}, \quad R_{P2} - R_{Pm1} = e_{Rp2}, \quad R_{P3} - R_{Pm1} = e_{Rp3}$$

Ou seja:  $R_{Pi} - R_{Pmj} = e_{Rpi}$ , onde  $i = 1, \dots, n$  e  $n =$  número de experimentos do conjunto de dados  
 $j = 1, \dots, m$  e  $m = 158$  séries

O mesmo procedimento foi usado para a razão de abundância  $R_W$ .

Portanto, as novas variáveis de trabalho definidas, serão o erro em  $R_P$ ,  $e_{Rp}$ , e o erro em  $R_W$ ,  $e_{Rw}$ , onde  $e_{Rpm}$  e  $e_{Rwm}$  são as médias do conjunto todo.

$$e_{Rpm} = \sum_{i=1}^n \frac{e_{Rpi}}{n}$$

$$e_{Rwm} = \sum_{i=1}^n \frac{e_{Rwi}}{n}$$

Após esta mudança de variáveis, o conjunto de dados experimentais assim constituído, é assumido como sendo uma amostra randômica, com variáveis normais randômicas  $e_{Rp}$  e  $e_{Rw}$ , onde  $e_{Rpm}$  e  $e_{Rwm}$  são as respectivas médias.

### 3.5 Análise Estatística dos Erros Experimentais

Os métodos clássicos de detecção de erros grosseiros assumem que as observações amostrais são de uma variável randômica distribuída normalmente. A essa amostra, aplica-se um teste estatístico apropriado para erros grosseiros sensível ao tipo de dispersão considerada, deduz-se a distribuição deste teste estatístico mediante a "hipótese nula" que todas as observações advêm da mesma população normal, e então rejeita-se a hipótese se o teste estatístico calculado for improvável de ter ocorrido em amostragem randômica. Mas a teoria e a prática de rejeição de erros grosseiros não são firmemente resolvidas, e se é feita em uma base puramente estatística é e sempre será um procedimento perigoso. Sua grande existência pode ser prova de que a população abaixo da linha é, em verdade, não o que se assumiu ser (Himmelblau, 1970):

Segundo Himmelblau (1970), dada uma amostra de observações  $X_1, X_2, \dots, X_n$  ( $n \geq 3$ ), a qual assume-se ser uma amostra randômica de uma variável randômica normal  $X$ , temos que o resíduo  $Y$  é  $Y = X_i - \bar{X}$ , onde  $\bar{X}$  é a média da variável  $X$ . Se o subscrito  $M$  é usado para designar a observação que tem resíduo maior, logo  $Y_M = X_M - \bar{X}$ .

Para um dado  $c$ , calculado pela equação 3.1, conduz-se o teste para  $X_M$ , e rejeita-se  $X_M$  se  $|Y_M| > cS_x$ , onde  $S_x$  é o desvio padrão da amostra. Do contrário, não se rejeita  $X_M$ . O valor de  $c$  pode mudar com o tamanho da amostra, e é dado explicitamente pela seguinte relação aproximada em termos da distribuição  $F$ :

$$c \approx \left( \frac{\nu}{n} \right)^{1/2} \left( \frac{3F_{1-\alpha}}{1 + [(3F_{1-\alpha} - 1)]^2 / (\nu + 1)} \right)^{1/2} \quad (3.1)$$

onde  $v = n-1$  (A raiz quadrada positiva é o valor de  $c$ ). O teste é conduzido usando-se a equação acima. Denota-se  $v/n$  por  $q$ , e encontra-se a porcentagem correspondente da proporção de variância,  $F_{1-q}$ , para 3 e para  $v$  graus de liberdade.

Seguindo este critério, foi feita então a análise estatística baseada no teste de hipóteses para médias, com um nível de confiança de 98% .

As variáveis usadas foram  $e_{Rp}$  e  $e_{Rw}$ , onde  $e_{Rpm}$  e  $e_{Rwm}$  são as médias e os resíduos  $Y_{Rp}$  e  $Y_{Rw}$ :

$$Y_{Rpi} = e_{Rpi} - e_{Rpm}$$

$$Y_{Rwi} = e_{Rwi} - e_{Rwm}$$

### 3.6 Aplicação de Rede Neural em ultracentrífugas

A aplicação de redes neurais na análise do desempenho separativo de ultracentrífugas se mostrou uma ferramenta muito poderosa (Migliavacca, 1999a, 1999b e 2000). Nesses estudos se verificou que o ajuste de dados com erros grosseiros apresenta uma dificuldade bem maior do que o ajuste de dados consistentes.

Dessa forma, este trabalho utiliza esta técnica de tratamento de dados para a identificação de erros grosseiros baseado no esquema de modelagem por redes neurais definido por Migliavacca (1999b) para um conjunto de dados semelhante, através do programa NEURO14. Este programa para treinamento de redes neurais (Nascimento, 1996), foi desenvolvido pelo Laboratório de Simulação e Controle de Processos do Departamento de Engenharia Química da USP em linguagem FORTRAN. De acordo com o programa, utiliza-se uma rede neural de três camadas, onde a informação segue em um único sentido ("feedforward neural network") e realiza o treinamento pelo algoritmo de retro-propagação. Para cada caso estudado foi definido o esquema das variáveis para o treinamento das redes neurais.

### 3.6.1 Primeiro caso: Rede Neural com poder de separação $\delta U$ como variável de saída

A primeira abordagem para detecção de erros grosseiros no primeiro conjunto de dados foi realizada, com base no esquema da Figura 3.2

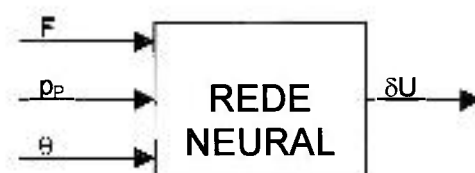


FIGURA 3. 2: Esquema das variáveis para treinamento da rede neural com  $\delta U$  como variável de saída

As variáveis de entrada e de saída inicialmente definidas foram:

entrada: -F : Fluxo de alimentação da ultracentrífuga (kg UF<sub>6</sub>/a)

-p<sub>p</sub> : Pressão na linha de retirada do produto (mbar)

- $\theta$  : corte : relação entre o fluxo da corrente enriquecida e o fluxo da alimentação .

saída: -  $\delta U$  : poder de separação de uma ultracentrífuga (variável que engloba os parâmetros separativos e os fluxos) (UTS/a).

Diferentemente do treinamento de redes para a modelagem do processo que divide o conjunto de dados experimentais em 2 conjuntos, “learning set” para o treinamento da rede propriamente dito, e “test set” para verificação do treinamento, no presente caso as redes são treinadas somente com “learning set”. A identificação dos pontos com possíveis erros requer que todos sejam avaliados da mesma forma, assim, o treinamento da rede deve ser realizado com um único conjunto de dados a ser apresentado.

Foram avaliadas redes com 6, 8 e 10 neurônios na camada interna. Essas redes foram então treinadas com até 50000 apresentações. Em cada um dos treinamentos foi realizada a comparação entre os valores experimentais do

poder de separação  $\delta U$  e os valores de  $\delta U$  calculados pela rede, através da diferença entre eles, resultando numa nova variável:

$$e_{N\delta U} = \delta U_{calc} - \delta U_{exp}$$

ou seja,  $e_{N\delta U}$  é a diferença entre o  $\delta U$  calculado pela rede neural e o  $\delta U$  experimental obtido no ensaio.

Foi utilizado o procedimento descrito no item 3.5 (Himmelblau, 1970), como critério de seleção de dados suspeitos de conterem erros grosseiros, porém, redefinindo as variáveis erro como  $e_{N\delta U}$ .

### 3.6.2 Segundo caso: Rede Neural com $R_P$ e $R_W$ (razões de abundância) como variáveis de saída

A segunda abordagem para detecção de erros grosseiros foi concebida com o intuito de permitir a aplicação direta da estimativa de erros realizada na análise estatística, e tratar as variáveis primárias diretamente, isto é, valores que no decorrer de um experimento são diretamente obtidos da instrumentação de processo. Assim as variáveis consideradas estão representadas no esquema da Figura 3.3.

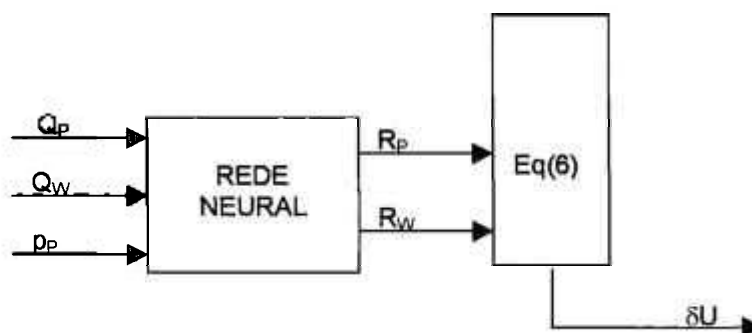


FIGURA 3. 3: Esquema proposto para as variáveis de treinamento da rede neural com  $R_P$  e  $R_W$  como variáveis de saída

As variáveis de entrada e de saída que foram agora definidas :

entrada:  $-Q_P$  : vazão de produto da ultracentrífuga (g/h)

$-Q_W$  : vazão de rejeito da ultracentrífuga (g/h)

$-p_P$  : Pressão na linha de retirada do produto (mbar)



saída: -  $R_p$  : razão de abundância do produto.

-  $R_w$  : razão de abundância do rejeito.

O tratamento por rede neural, com a estrutura de rede montada segundo o esquema de variáveis apresentado na figura 3.3, permitiu aproveitar o resultado do procedimento da análise estatística de  $e_{Rp}$  e  $e_{Rw}$ , descrito no item 3.5, segundo Himmelblau (1970), como critério de seleção de dados suspeitos de conterem erros grosseiros. Porém, neste caso, as variáveis erro foram redefinidas como:

$$e_{NRp} = R_p \text{ calc} - R_p \text{ exp}$$

ou seja,  $e_{NRp}$  é a diferença entre o  $R_p$  calculado pela rede neural e o  $R_p$  experimental obtido no ensaio.

Portanto, a diferença entre a média dos  $e_{NRp}$  e cada  $e_{NRp}$  é o resíduo  $Y_{NRp}$ .

Analogamente, para  $R_w$  tem-se:

$$e_{NRw} = R_w \text{ calc} - R_w \text{ exp}, \text{ sendo que o resíduo é } Y_{NRw}.$$

Neste caso foram treinadas redes com número de neurônios variando de 6 a 10 e número de apresentações variando de 1000 a 50000 apresentações.

A escolha da melhor rede foi feita avaliando-se o decréscimo da curva do erro total durante o treinamento da rede. O critério heurístico adotado foi de escolher um número de apresentações mínimo para cada rede a partir do qual o erro global decresce pouco com a continuação do treinamento. Para valores muito altos do número de apresentações, pode ocorrer um indesejável ajuste de dados com erros grosseiros ao modelo.

## 4 RESULTADOS

### 4.1 Balanço Material Macroscópico

Dos 791 experimentos realizados, os quais compõem o primeiro conjunto de dados original, 27 apresentaram condições insuficientes para realização de análise por espectrometria de massa da amostra, e estão apresentados na Tabela 4.1.

TABELA 4. 1 : Relação dos experimentos rejeitados por insuficiência de material para análise por espectrometria de massa

55	450	630	647	654	739
439	460	631	648	655	761
441	461	632	651	656	
442	462	633	652	663	
446	629	634	653	664	

Portanto foi necessário desconsiderá-los antes mesmo de uma análise detalhada de dados, resultando em um conjunto de 764 experimentos. Porém, os números dos experimentos foram preservados na ordem de execução destes dentro do ensaio.

A aplicação do procedimento do balanço material macroscópico no conjunto de dados com 764 experimentos, apontou 176 com erros grosseiros, indicados na Tabela 4.2. Esses 176 experimentos são então excluídos do conjunto de dados original, pois não obedecem ao balanço de massa. Desta forma, o conjunto de dados experimentais passa a ter 588 experimentos. Esse é o conjunto onde são aplicadas as diferentes técnicas para detecção de erros, dentre elas a detecção através de redes neurais.

TABELA 4. 2 : Relação dos experimentos que não respeitam o balanço material macroscópico

50	203	261	301	343	419	609	690	752
58	204	262	302	348	420	610	691	757
79	205	278	303	349	421	614	698	759
98	206	279	304	350	422	615	701	760
99	207	280	305	351	425	616	707	765
100	208	281	306	352	426	649	709	766
107	212	282	307	362	427	650	710	767
108	213	283	313	372	437	657	714	769
131	214	284	320	373	438	658	716	770
176	227	286	321	374	440	659	719	771
177	228	290	322	375	458	660	720	777
178	229	291	323	383	466	661	724	784
179	230	292	324	399	467	662	725	786
180	231	293	325	400	468	668	727	789
181	232	295	326	401	470	669	733	790
194	254	296	327	402	478	670	734	791
195	255	297	328	403	490	671	735	
196	256	298	337	404	491	672	736	
198	257	299	341	408	492	673	750	
199	259	300	342	409	495	677	751	

## 4.2 Análise de “Cluster”

O conjunto de dados com 588 experimentos foi submetido a uma análise por “clusters”. A análise contemplou a avaliação das respostas obtidas em termos do poder de separação  $\delta U$  de cada experimento, por ser esta a variável mais significativa e mais sensível na análise de desempenho separativo de uma ultracentrífuga.

As variáveis utilizadas foram o fluxo de alimentação  $F$  em 8 níveis, o corte  $\theta$  em 7 níveis e a pressão na retirada do produto  $p_p$  em 5 níveis. Quando um resultado diferente do poder de separação  $\delta U$  é observado em uma série de dados similares, o experimento correspondente é indicado como suspeito. Foram detectados pontos que estavam dispersos em relação às linhas de tendência, pois o comportamento geral do poder de separação  $\delta U$ , em função do fluxo de

alimentação, do corte  $\theta$  ou da pressão na linha de produto  $p_p$ , mantidas constantes duas destas variáveis, é conhecido e foi apresentado em Migliavacca, 1999a, 1999b e 2000. Estes cinco experimentos detectados através da análise de “cluster” são apresentados na Tabela 4.3:

TABELA 4. 3 : Experimentos considerados suspeitos pela análise de cluster

Nº dos Experimentos	
415	689
423	732
537	

### 4.3 Análise Estatística dos Erros Experimentais utilizando Teste de Hipóteses

Nas Figuras 4.1 e 4.2 são mostradas as dispersões dos resíduos das variáveis de estudo  $Y_{Rp}$  e  $Y_{Rw}$ , para os 588 experimentos correspondentes.

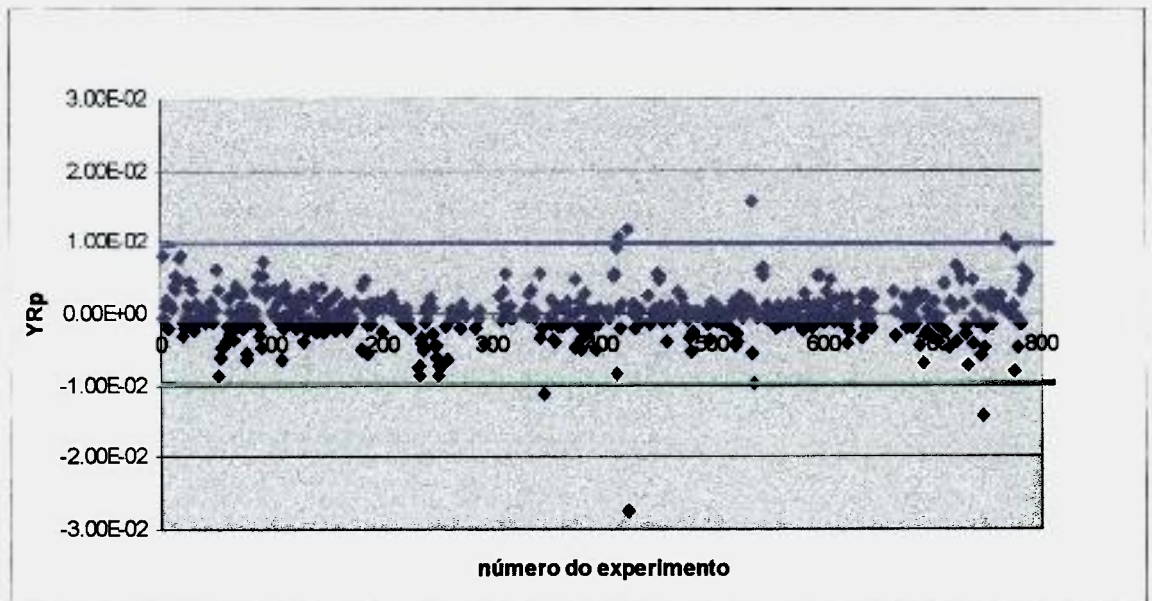


FIGURA 4. 1: Dispersão dos resíduos  $Y_{Rp}$  com a seqüência dos experimentos

Segundo o procedimento descrito na seção 3.5, foi realizado o teste de hipóteses dos erros experimentais e os parâmetros calculados pela análise estatística estão apresentados na Tabela 4.4.

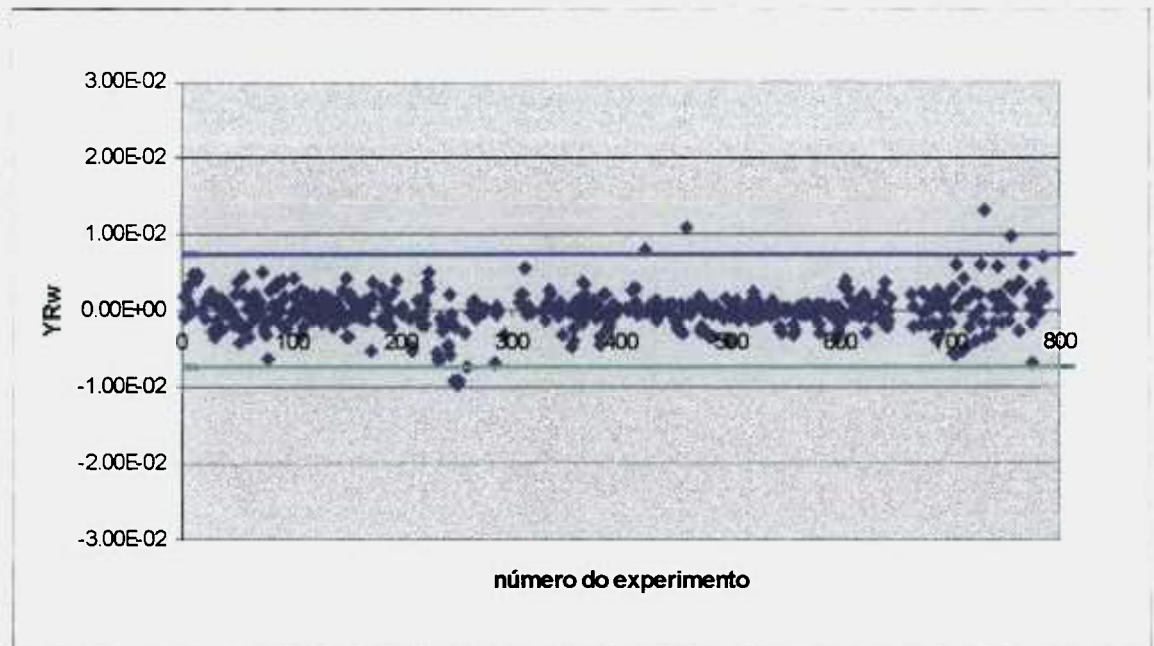


FIGURA 4. 2: Dispersão dos resíduos  $Y_{Rw}$  com a seqüência dos experimentos

TABELA 4. 4 : Teste de Hipóteses dos erros experimentais

Nº de exptos	239	
<b>C</b>	3,100919	
<b>Variável</b>	$\epsilon_{RP}$	$\epsilon_{RW}$
Mínimo	-0,0143	-0,01022
Máximo	0,01175	0,00871
Média	-0,001195	-0,00656
Desvio Padrão S	0,004680	0,003247
Média + c*S	0,013325	0,009412
Média - c*S	-0,015715	-0,010724

Assim, foram detectados pontos, indicados na Tabela 4.5, que estão fora do intervalo definido pela análise da distribuição, ou seja, dados experimentais que são considerados suspeitos de conterem erros por não se ajustarem muito bem em torno da média, dentro do limite definido pela análise estatística.

TABELA 4. 5 : Relação dos experimentos considerados suspeitos pela análise estatística

CRP	CRW	CRP + CRW
347	251	251
415	252	252
423	253	253
424	248	248
768	260	260
537	423	347
746	459	415
	732	423
	756	424
	768	459
		537
		732
		746
		756
		768

#### 4.4 Primeiro caso: Rede Neural com poder de separação $\delta U$ como variável de saída

Foram testadas redes com 6, 8 e 10 neurônios na camada oculta com até 50000 apresentações para o treinamento da rede. A Tabela 4.6 apresenta os valores dos erros globais mínimos (dados pela somatória das diferenças entre os valores experimentais e os calculados pela rede de todo o conjunto de dados) encontrados quando permitidas até 50000 apresentações. Por meio desses valores, temos a indicação de que a rede que seria a melhor candidata à modelagem dos dados experimentais seria a de 10 neurônios.

TABELA 4. 6 : Erro global mínimo encontrado para cada rede neural aplicada permitidas até 50000 apresentações no treinamento

No. de neurônios na camada interna	Erro global mínimo	No. de apresentações para o erro mínimo
6	0,71817	2100
8	0,85954	50000
10	0,70677	28000

O desenvolvimento dos erros globais ao longo do treinamento dessas redes é mostrado nas Figuras 4.3 e 4.4, onde se observa que após 5000 apresentações o erro decai lentamente com a continuidade do treinamento das redes. Para a rede de 6 neurônios o erro global mínimo é atingido com 2100 apresentações.

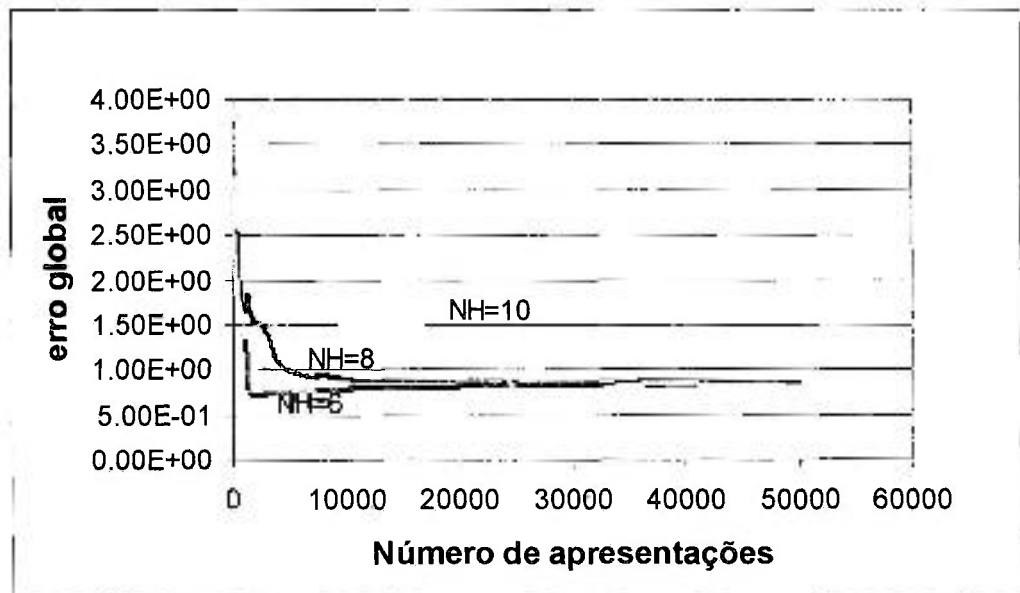


FIGURA 4. 3: Erro global das redes com 6, 8 e 10 neurônios com treinamento total até 50000 apresentações - variável de saída:  $\delta U$

A rede de 10 neurônios apresenta o menor erro total mínimo, porém isto ocorre com um número de apresentações muito elevado, e para a identificação de erros grosseiros, esse refinamento no treinamento não é interessante.

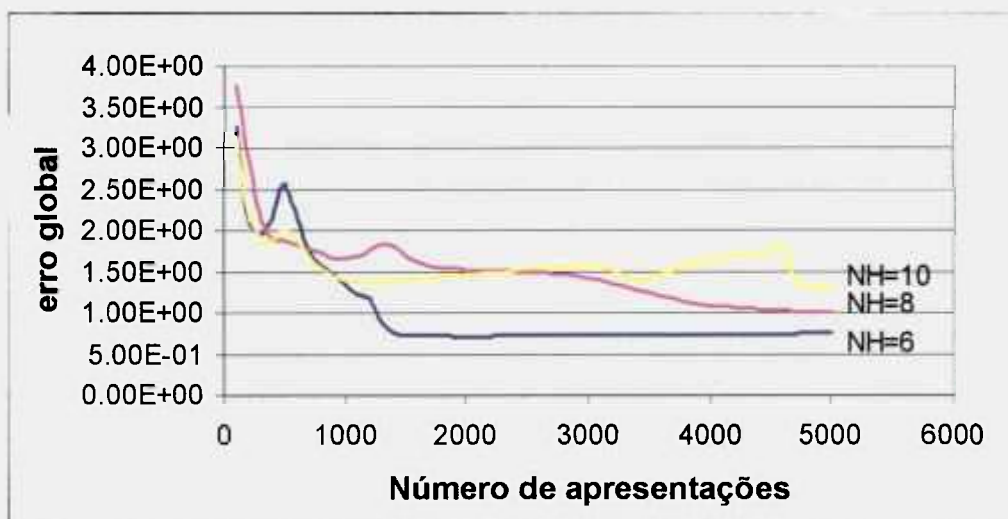


FIGURA 4. 4: Erro global das redes com 6, 8 e 10 neurônios com treinamento total até 5000 apresentações - variável de saída:  $\delta U$

Nota-se que, até 5000 apresentações, as redes de 6 e de 8 neurônios apresentam melhores curvas de decaimento do erro global. Portanto, estas redes foram escolhidas para o processo de detecção de erros neste conjunto de dados e para esta configuração de rede. Percebe-se que com um número pequeno de apresentações as redes atingem valores mínimos razoáveis, sendo que a rede de 6 neurônios possui um erro global mínimo para 5000 apresentações menor que a de 8.

Nas Figuras 4.5 e 4.6, observa-se a comparação entre os valores experimentais e os calculados das redes de 6 e de 8 neurônios na camada oculta treinada com 5000 apresentações. Os pontos que se distanciam da nuvem ao redor da diagonal são então considerados dados suspeitos de conterem erros grosseiros.



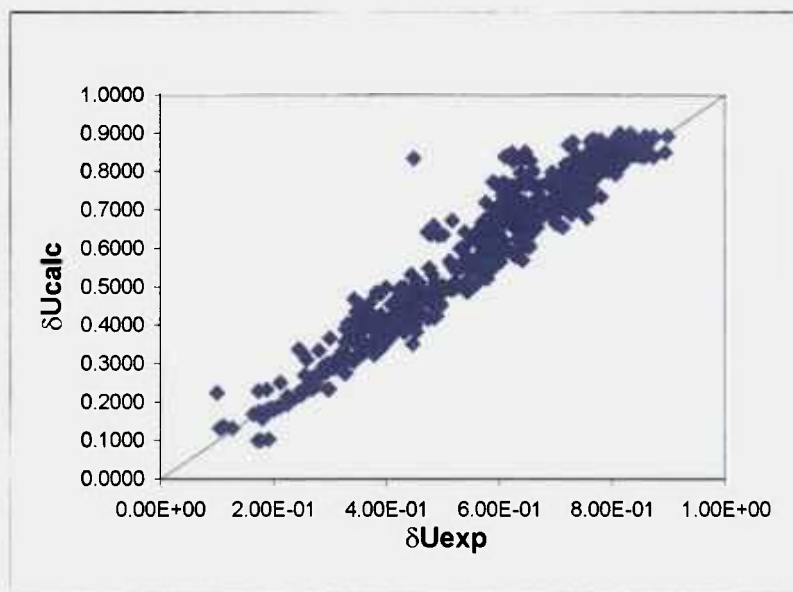


FIGURA 4. 5: Comparação entre os valores experimentais e calculados do poder de separação  $\delta U$  pela rede neural de 6 neurônios na camada oculta com 5000 apresentações

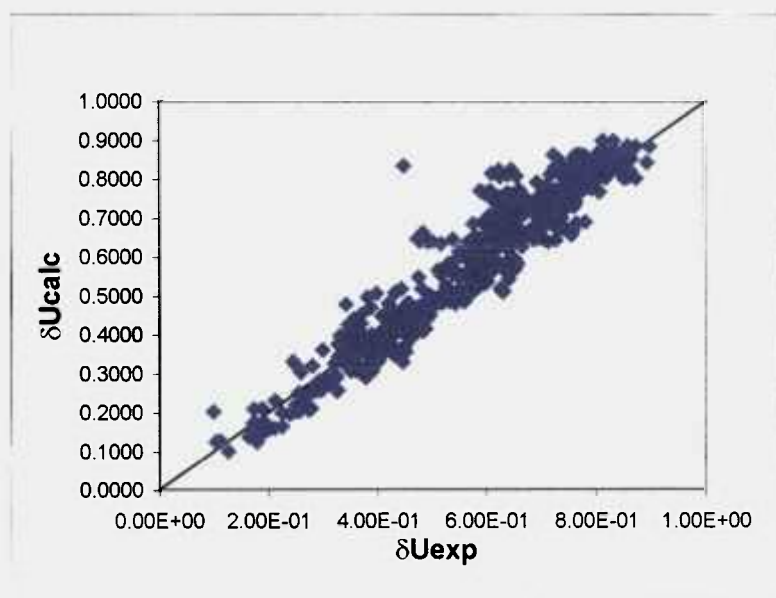


FIGURA 4. 6: Comparação entre os valores experimentais e calculados do poder de separação  $\delta U$  pela rede neural de 8 neurônios na camada oculta com 5000 apresentações

Através do procedimento descrito no item 3.5 (Himmelblau, 1970), foi efetuada a identificação de dados suspeitos de conterem erros grosseiros. Os experimentos apontados são então retirados do conjunto de dados, o treinamento de rede neural é refeito, e o cálculo do teste de hipóteses é novamente aplicado para busca de novos erros grosseiros.

Quando novos suspeitos são apontados, o processo de exclusão e nova rodada foram repetidos até que não houvesse mais experimentos suspeitos apontados. As Tabelas 4.7 e 4.8 mostram o resumo das identificações de experimentos suspeitos de conterem erros grosseiros, para a rede de 6 neurônios com 5000 e 1500 apresentações, e para a rede de 8 neurônios com 5000 apresentações, respectivamente. Além disso, nas próprias tabelas são apresentados os testes de hipóteses para o conjunto de dados, após o treinamento das respectivas redes. O número zero indica que a rede não detectou mais nenhum dado suspeito.

TABELA 4. 7 : Teste de Hipóteses dos resultados das redes neurais com 6 neurônios na camada oculta

Nº apresentações	5000		1500	
	1	2	1	2
rodada				
Nº experimentos	588	585	588	585
c	3,122015	3,121941	3,122015	3,121941
mínimo	-1,009500	-1,595500	-1,047800	-1,547300
máximo	2,294900	1,545900	2,296000	1,184900
média	-0,135747	-0,167859	-0,135902	-0,169502
desvpad S	0,357477	0,612401	0,352797	0,497530
média + c*S	0,980303	1,744020	0,965535	1,383757
média - c*S	-1,251796	-2,079739	-1,237338	-1,722761
experimentos	423	0	423	0
suspeitos	635		635	
identificados	636		636	

TABELA 4. 8 : Teste de Hipóteses dos resultados da rede neural com 8 neurônios na camada oculta para 5000 apresentações

rodada	1	2	3	4	5
Nº experimentos	588	586	583	581	579
c	3,122015	3,121966	3,121891	3,121840	3,121789
mínimo	-1,096400	-1,086400	-1,129800	-1,137700	-1,031900
máximo	3,368000	1,028400	0,912700	0,8639006	0,733200
média	-0,184229	-0,175834	-0,187418	-0,204336	-0,213890
desvpad S	0,383508	0,354790	0,344559	0,339902	0,316910
média + c*S	1,013090	0,931810	0,888256	0,856784	0,775437
média - c*S	-1,381548	-1,283477	-1,263092	-1,265456	-1,203217
experimentos	423	635	612	637	0
suspeitos	636	641	746	732	
identificados		744			

#### 4.5 Segundo caso: Rede Neural com $R_p$ e $R_w$ (razões de abundância) como variáveis de saída

Neste caso, foram treinadas redes com número de neurônios variando de 6 a 10 e número de apresentações variando de 1000 a 50000 apresentações.

A escolha da rede que melhor se adequasse ao processo de detecção de erros grosseiros para esse conjunto de dados, foi feita avaliando-se a curva de decréscimo do RMST (valor do erro total de treinamento da rede), tomando cuidado para não provocar superajuste, no caso de um número alto de apresentações. Na Figura 4.7, são mostradas as curvas do erro total de treinamento (RMST) para as três estruturas de rede avaliadas ( $NH=6, 8$  e  $10$ ) permitidas até 5000 apresentações. Nestas condições, a rede de 8 neurônios na camada oculta apresentou o menor valor de erro total com um mínimo de apresentações após o decaimento inicial. O ganho obtido a partir de 5000 apresentações não foi muito grande com relação ao erro de ajuste do treinamento. Na realidade, para valores altos do número de apresentações, ocorre um indesejável ajuste dos erros grosseiros ao modelo, o que não é o objetivo deste trabalho.

Portanto a configuração de rede escolhida foi de 8 neurônios e até 5000 apresentações.

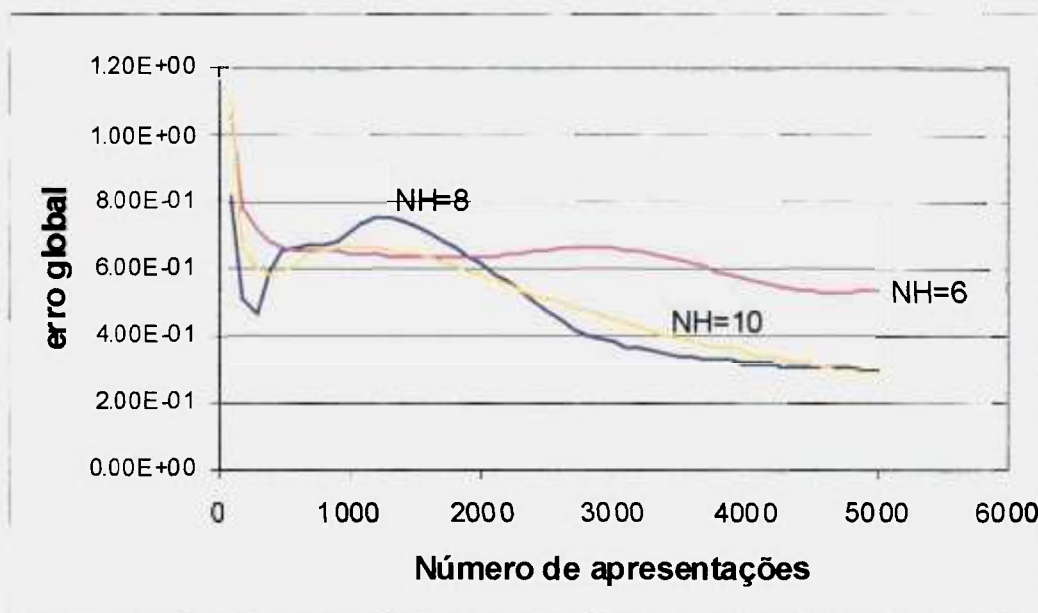
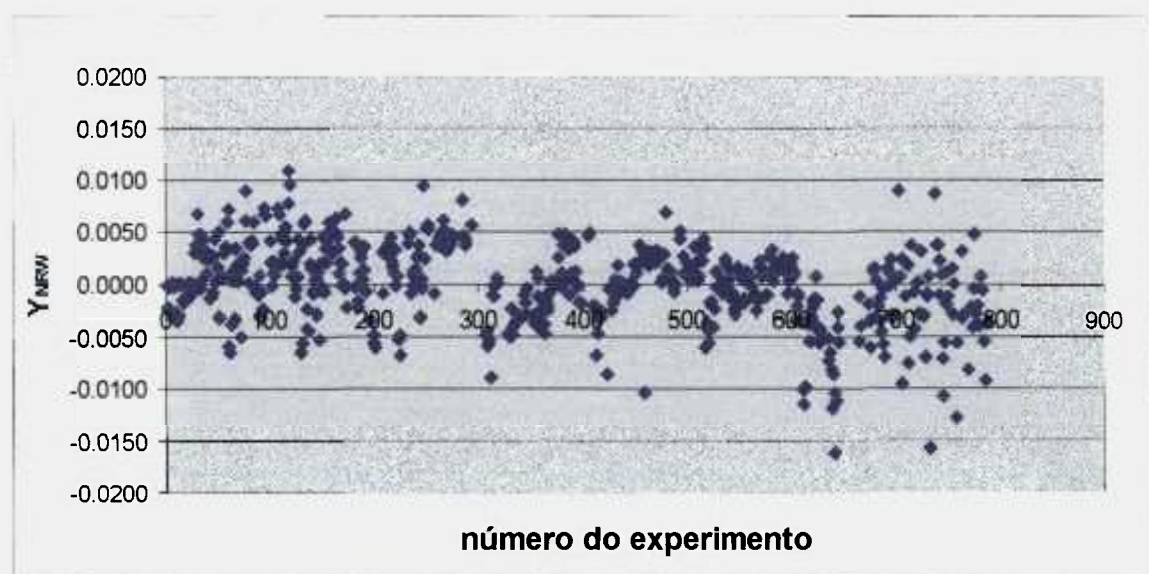


FIGURA 4. 7: Erro global das redes com 6, 8 e 10 neurônios com treinamento total até 5000 apresentações - variáveis de saída:  $R_P$  e  $R_W$

Assim definida a rede, fez-se o treinamento desta com os 588 experimentos. Os resultados foram analisados através da diferença entre o valor calculado pela rede e o experimental para as duas variáveis de saída  $R_P$  e  $R_W$ . Estas diferenças foram denominadas de  $e_{NR_P}$  para a variável  $R_P$  e  $e_{NR_W}$  para a variável  $R_W$ . Foram calculados os resíduos  $Y_{NR_P}$  e  $Y_{NR_W}$  das variáveis  $e_{NR_P}$  e  $e_{NR_W}$  em relação às suas respectivas médias; estes resíduos estão apresentados nas Figuras 4.8 e 4.9 em ordem seqüencial.

Analisando estes gráficos, observa-se que durante o ensaio existiram características diferentes entre alguns grupos de experimentos, implicando na definição de 3 regiões distintas. Durante a execução do ensaio, o qual levou alguns meses, ocorreu uma progressiva degradação da calibração de alguns instrumentos (sensores de vazão), além do fato de que a temperatura também variou durante o período de realização do ensaio. E é justamente nas mudanças de comportamento dos experimentos observadas nas Figuras 4.8 e 4.9, onde se distingue as três fases ocorridas ao longo do ensaio

FIGURA 4. 8: Análise dos resíduos para a variável de saída  $R_p$ FIGURA 4. 9: Análise dos resíduos para a variável de saída  $R_w$ 

Próximo aos experimentos codificados de 294 a 308, foi detectado problema através da verificação do balanço de massa descrito no item 3.2., e então, os instrumentos foram recalibrados. O mesmo problema ocorreu antes do experimento de número 604. Este fato foi apontado pela rede neural.

Portanto, pode-se definir três regiões distintas, formando três subconjuntos distintos, a saber:



Primeiro: do experimento 1 até o 294, resultando em 239 experimentos.

Segundo: do experimento 308 até o 604, resultando em 236 experimentos.

Terceiro: do experimento 605 até o 788, resultando em 113 experimentos.

Desse modo, as redes foram treinadas com 3 conjuntos de dados, sendo uma rede independente da outra. As configurações das redes foram definidas baseada na anterior para o conjunto inteiro de dados (588 experimentos). Após alguns treinamentos, verificou-se que, também para os 3 grupos ou subconjuntos de dados, a melhor configuração de rede era de 8 neurônios na camada oculta. Quanto ao número de apresentações, decidiu-se manter o mesmo de 5000, como padronização para os 3 subconjuntos, como é mostrado nas Figuras 4.10, 4.11 e 4.12.

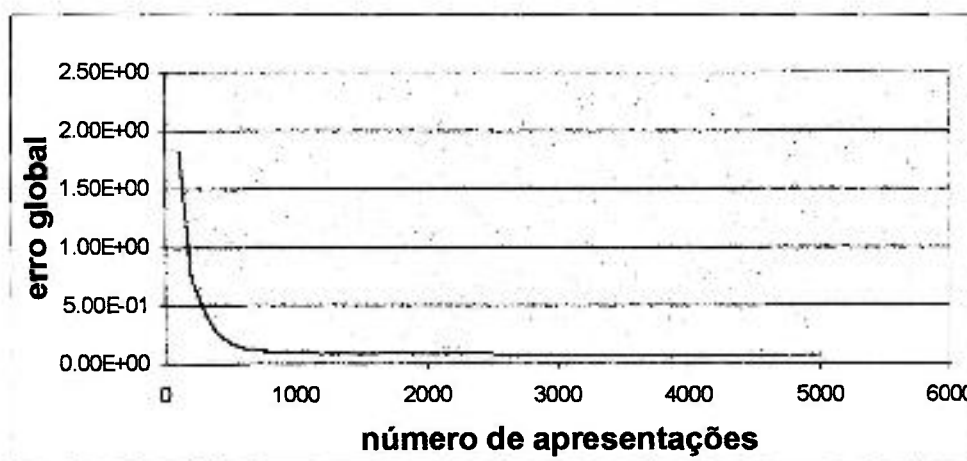


FIGURA 4. 10: Erro global no treinamento da rede neural com 8 neurônios na camada oculta para o primeiro subconjunto

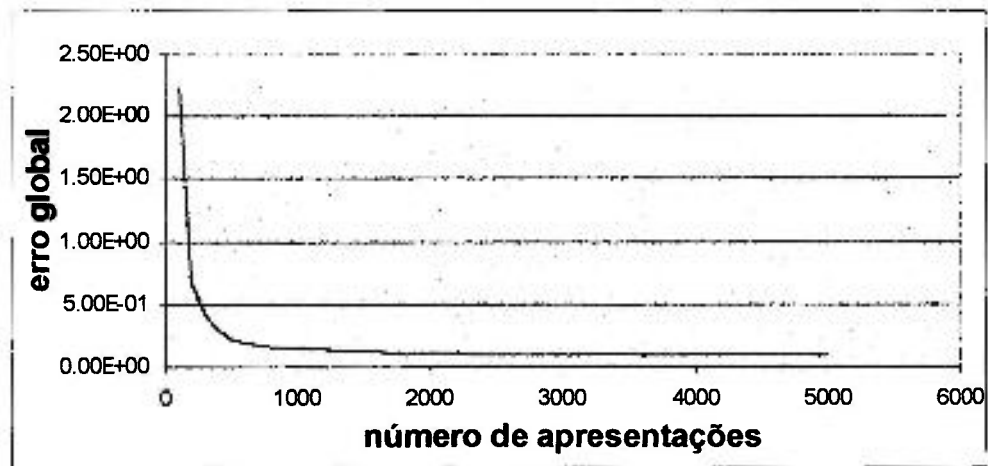


FIGURA 4. 11: Erro global no treinamento da rede neural com 8 neurônios na camada oculta para o segundo subconjunto

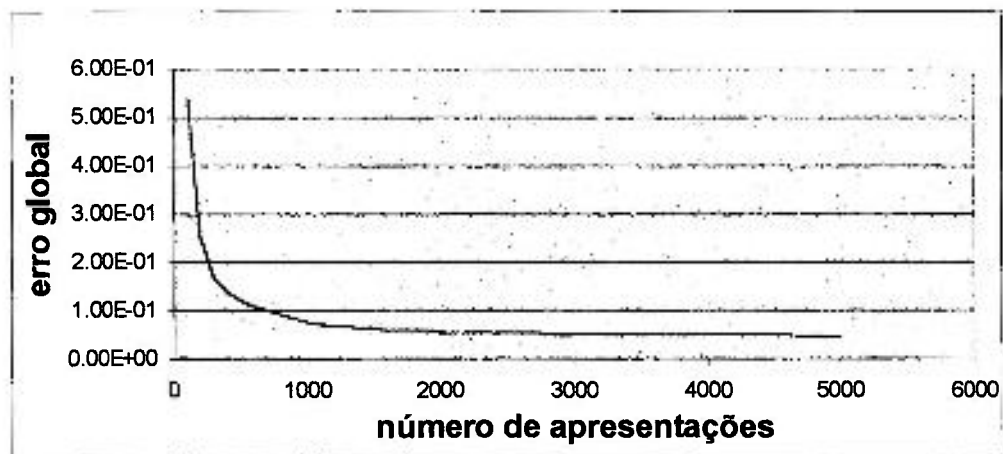


FIGURA 4. 12: Erro global no treinamento da rede neural com 8 neurônios na camada oculta para o terceiro subconjunto

Em cada um dos treinamentos de cada grupo foi realizada a análise de resíduos para as variáveis  $e_{NRp}$  e  $e_{NRw}$ , que são os erros relativos as diferenças entre os valores calculados pelas redes e os experimentais de  $R_p$  e  $R_w$ , respectivamente, como mostrado nas figuras 4.13 a 4.18.

Como critério de detecção de erros grosseiros foi escolhido o critério definido por Himmeublau (1970), já descrito anteriormente no item 3.5 deste trabalho.

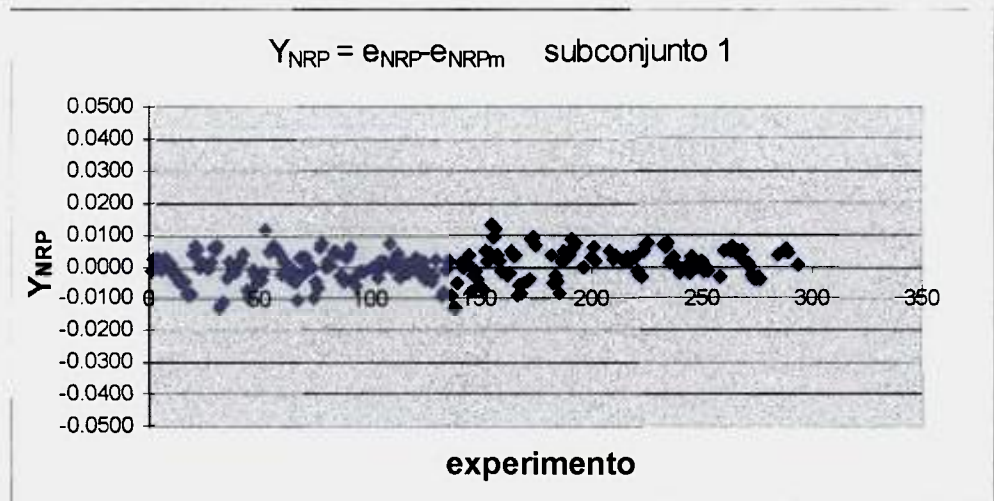


FIGURA 4. 13: Análise dos resíduos para a variável de saída  $R_P$  para o 1º subconjunto

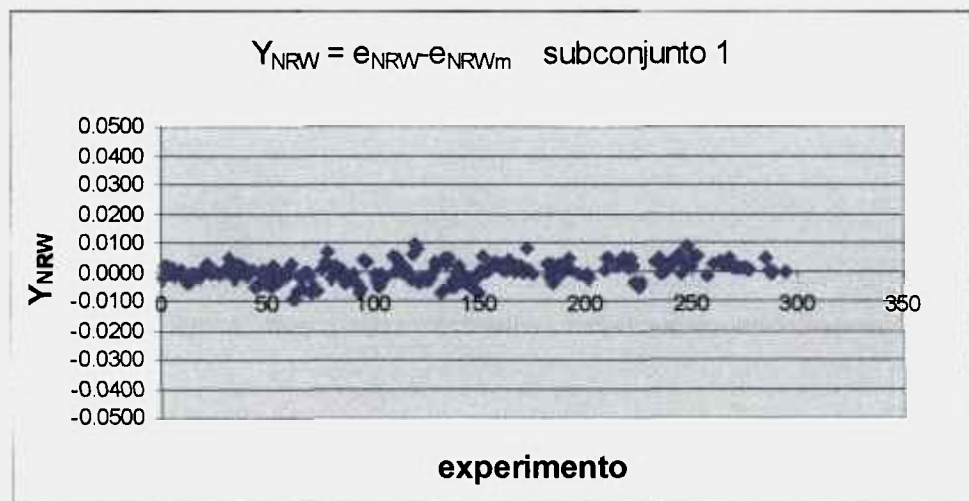


FIGURA 4. 14: Análise dos resíduos para a variável de saída  $R_W$  para o 1º subconjunto



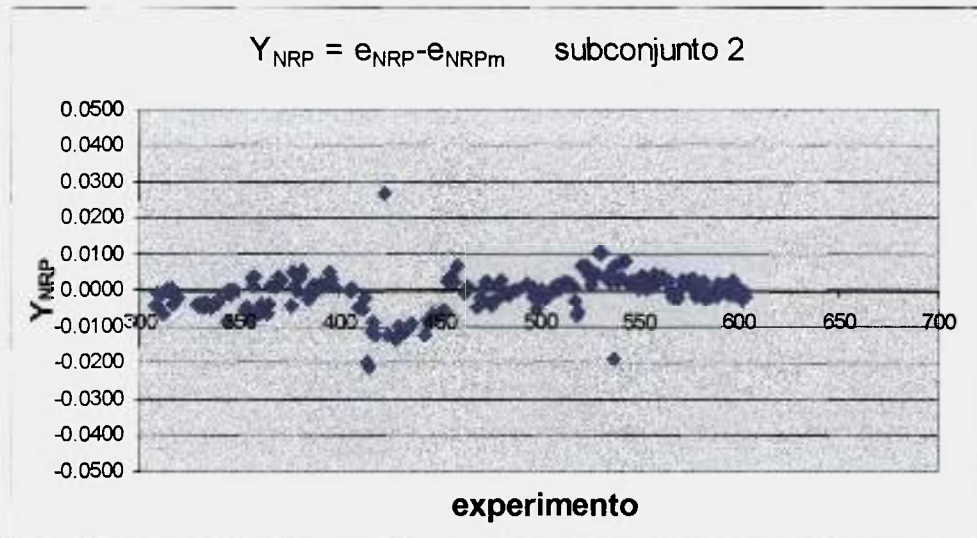


FIGURA 4. 15: Análise dos resíduos para a variável de saída  $R_P$  para o 2º subconjunto

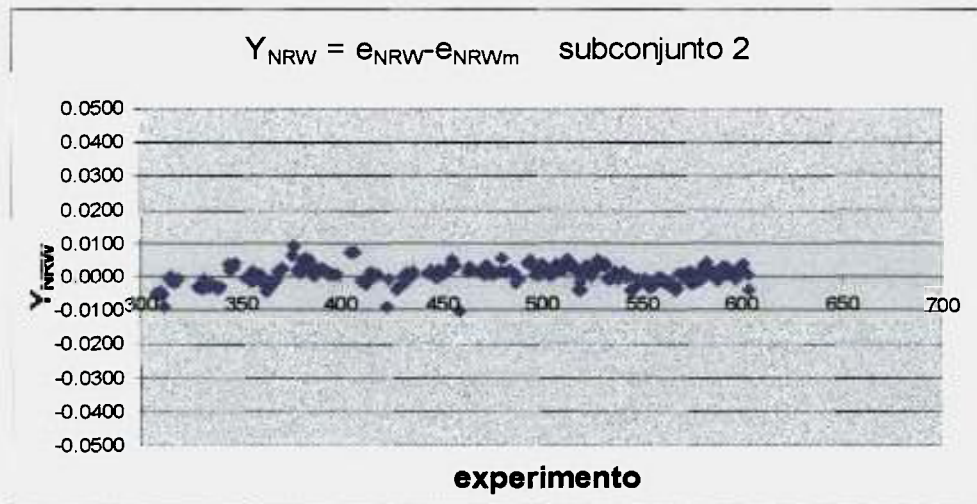


FIGURA 4. 16: Análise dos resíduos para a variável de saída  $R_W$  para o 2º subconjunto

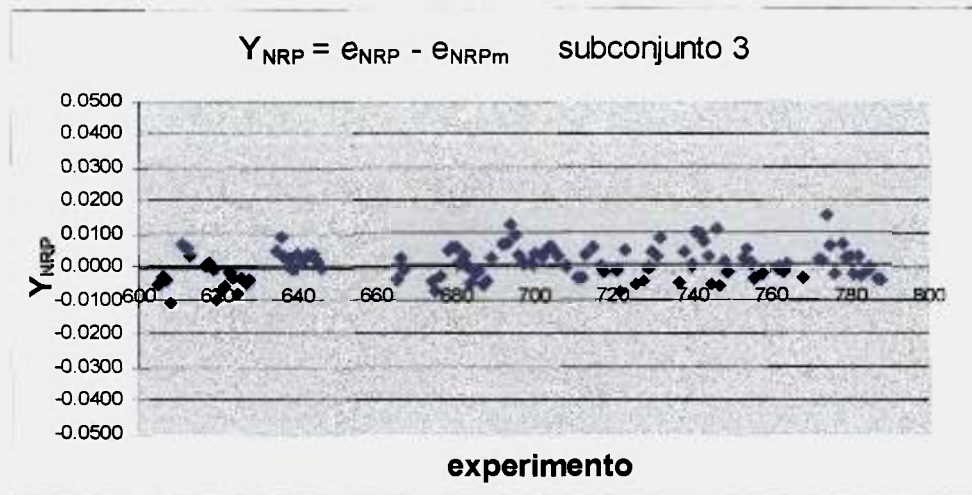


FIGURA 4. 17: Análise dos resíduos para a variável de saída  $R_P$  para o 3º subconjunto

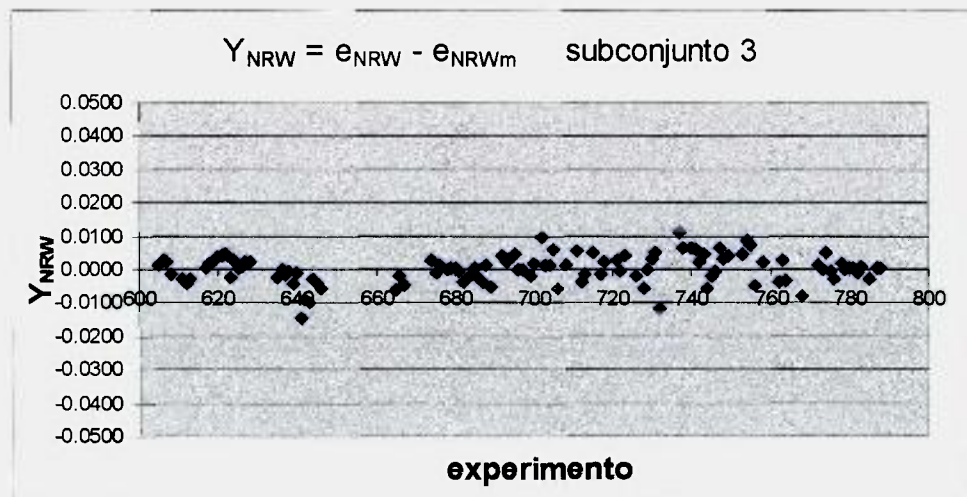


FIGURA 4. 18: Análise dos resíduos para a variável de saída  $R_W$  para o 3º subconjunto

Após cada treinamento de rede para cada um dos subconjuntos, foi calculado o valor de  $c$  segundo procedimento descrito por Himmelblau (1970).

Nas Tabelas 4.9, 4.10 e 4.11 são apresentados os testes de hipóteses para os três subconjuntos de dados, após o treinamento das redes com 8 neurônios na camada oculta.

TABELA 4. 9 : Teste de Hipóteses dos resultados das redes neurais com 8 neurônios na camada oculta para o 1° subconjunto de dados

rodada	1	
N° experimentos	239	
c	3,100919	
variável	$e_{NRP}$	$e_{NRW}$
mínimo	-0,0143	-0,01022
máximo	0,01175	0,00871
média	-0,001195	-0,00656
desvpad S	0,004680	0,003247
média + c*S	0,013325	0,009412
média - c*S	-0,015715	-0,010724
experimentos suspeitos identificados	0	0

TABELA 4. 10 : Teste de Hipóteses dos resultados das redes neurais com 8 neurônios na camada oculta para o 2° subconjunto de dados

rodada	1		2		3	
N° experimentos	236		229		227	
c	3,100468		3,099366		3,099048	
variável	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$
mínimo	-0,02127	-0,01051	-0,014030	-0,006250	-0,015760	-0,010910
máximo	0,02701	0,0089	0,009900	0,007700	0,009920	0,006160
média	-0,000581	0,000354	-0,000654	0,000010	-0,001020	-0,000495
desvpad S	0,005261	0,002743	0,004307	0,002246	0,004471	0,003220
média + c*S	0,015730	0,008857	0,012695	0,006973	0,012835	0,009485
média - c*S	-0,016893	-0,008149	-0,014002	-0,006952	-0,014876	-0,010475
experimentos	414	312	428	376	418	311
suspeitos	415	377				
identificados	423	423				
	537	459				
rodada	4		5			
N° experimentos	225		223			
c	3,098708		3,098369			
variável	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$		
mínimo	-0,016300	-0,010830	-0,014420	-0,008910		
máximo	0,010140	0,006050	0,009410	0,005710		
média	-0,000981	-0,000604	-0,001021	-0,000560		
desvpad S	0,004557	0,003263	0,004597	0,002862		
média + c*S	0,013140	0,009507	0,013221	0,008308		
média - c*S	-0,015102	-0,010714	-0,015262	-0,009428		
experimentos	416	309	0	0		
suspeitos						
identificados						

TABELA 4. 11 : Teste de Hipóteses dos resultados das redes neurais com 8 neurónios na camada oculta para o 3º subconjunto de dados

rodada	1		2	
Nº experimentos	113		112	
c	3,061339		3,060669	
variável	$\epsilon_{NRP}$	$\epsilon_{NRW}$	$\epsilon_{NRP}$	$\epsilon_{NRW}$
mínimo	-0,01102	-0,01463	-0,011050	-0,011790
máximo	0,01534	0,01132	0,015320	0,011340
média	0,000643	0,000192	0,000748	0,000169
desvpad S	0,004835	0,004235	0,004742	0,003507
média + c*S	0,015444	0,013155	0,015260	0,010902
média - c*S	-0,014158	-0,012771	-0,013764	-0,010564
experimentos suspeitos identificados	0	641	0	0

Através deste critério, dados suspeitos de conterem erros grosseiros foram detectados. Para cada treinamento de rede neural foi feito o cálculo do teste de hipóteses, obtendo assim, após cada rodada, os experimentos considerados suspeitos.

Os experimentos indicados pelas redes foram retirados dos subconjuntos e estes submetidos a um novo treinamento de rede para ser avaliada a presença de novos experimentos suspeitos.

A cada nova rodada de treinamentos e detecção de suspeitos nos conduziu à relação de experimentos suspeitos de conterem erro grosseiro apresentada na Tabela 4.12.

Os treinamentos de redes foram feitos até não se identificar mais a presença de erros grosseiros, sempre aplicando o critério de detecção citado no item 3.5.

TABELA 4. 12 : Indicação de dados suspeitos de conterem erros grosseiros pelas redes neurais para os 3 subconjuntos de dados após cada rodada

Rodada	1º subconjunto	2º subconjunto	3º subconjunto
1	0	312 423 377 459 414 537 415	641
2		376 428	0
3		311 418	
4		309 416	
5		0	

## **5 DISCUSSÃO DOS RESULTADOS**

### **5.1 Balanço material macroscópico, análise de cluster e análise estatística**

Do conjunto de dados original com 791 experimentos foram excluídos 27 experimentos com amostras insuficientes e 176 experimentos que não respeitaram o balanço material, como apresentado nas Tabelas 4.1 e 4.2, respectivamente.

Esse conjunto com 588 experimentos foi submetido a uma análise de “cluster” que indicou 5 experimentos como suspeitos de conterem erros grosseiros apresentados na Tabela 4.3, enquanto a análise estatística dos erros da razões de abundância apontou 15 experimentos mostrados na Tabela 4.5. Quatro desses experimentos são apontados por ambos os critérios. Assim, a união desses 2 conjuntos de suspeitos, obtidos pela análise estatística e pela análise de clusters, nos fornece um total de 16 experimentos indicados como experimentos suspeitos de conterem erros grosseiros.

Alguns procedimentos estatísticos convencionais foram testados e aplicados ao primeiro conjunto de dados, na tentativa de se encontrar o melhor critério estatístico de detecção de valores suspeitos de conterem erros grosseiros através do intervalo entre quartis modificado, (Lapponi,2000), e o método de detecção usando t-student considerando um nível de confiança de 95%. Ambos se mostraram rigorosos demais na identificação dos experimentos suspeitos. O conjunto de dados pode conter erros estocásticos e sistemáticos, que não necessariamente sejam considerados como erros grosseiros, só por não estarem dentro do intervalo de confiança estabelecido.

Logo, o procedimento descrito por Himmelblau (1970), e apresentado no item 3.5, se mostrou menos rigoroso e ao mesmo tempo adequado para identificação dos experimentos suspeitos de conterem erros grosseiros.

## 5.2 Primeiro caso: Rede Neural com poder de separação $\delta U$ como variável de saída

Na tabela 4.6 estão apresentados os valores dos erros totais mínimos obtidos pelas estruturas de redes testadas quando permitidas até 50000 apresentações. A rede com 10 neurônios possui o menor valor para o erro total mínimo. Entretanto, observando as Figuras 4.3 e 4.4, nota-se que para poucas apresentações, a rede de 10 neurônios apresenta ainda um erro total muito grande em relação às demais redes. Após a acomodação inicial dos pesos, o número de apresentações deve ser a menor possível, para que a rede não cause à adaptação dos experimentos contendo erros grosseiros. A rede de 10 neurônios não possui esta característica.

No caso da rede de 6 neurônios, esta possui um valor mínimo para o erro total do treinamento apresentado na Tabela 4.6. Nota-se, porém, que após a retirada dos experimentos suspeitos identificados, o desvio padrão aumenta de valor em relação ao de antes da retirada. Isto indica que o ajuste da rede proposta, ao novo conjunto de dados, já não satisfaz como antes.

Entretanto, através dos resultados apresentados na Tabela 4.8, pode-se constatar que o desvio padrão da rede de 8 neurônios com 5000 apresentações varia pouco de uma rodada para outra, e a retirada dos erros grosseiros faz com que esse desvio diminua, indicando um melhor ajuste da rede aos dados experimentais. Portanto, decidiu-se escolher a rede de 8 neurônios com 5000 apresentações como ferramenta de detecção dos erros grosseiros para conjunto de dados estudado, nessa configuração de rede que tem o  $\delta U$  como variável de saída.

Na Tabela 4.7, são apresentados os três experimentos identificados pela rede de 6 neurônios como contendo erros grosseiros. Dentre estes experimentos, somente um foi detectado pela análise estatística e de cluster, mostrados na Tabela 4.5 e 4.3, respectivamente.

Com relação à rede de 8 neurônios, dos nove experimentos identificados por ela, que constam da Tabela 4.8, três foram detectados pela análise estatística e de "cluster". Os outros seis experimentos detectados pela



rede, que representam cerca de 66% dos experimentos apontados, não foram identificados anteriormente. Estes experimentos foram verificados e constatou-se que foram influenciados pela mudança de fluxo ou de pressão, e conseqüentemente uma amostragem realizada em regime transiente de processo. A análise estatística identificou dados que não foram apontados pela rede neural, mostrando que o rigorismo estatístico foi maior que o desejado, frente à qualidade dos dados experimentais.

### **5.3 Segundo caso: Rede Neural com $R_p$ e $R_w$ (razões de abundância) como variáveis de saída**

Segundo análise feita através dos gráficos das Figuras 4.8 e 4.9, constatou-se que durante o ensaio existiram características diferentes entre alguns grupos de experimentos. Como o ensaio teve uma duração muito longa, ocorreram alterações tanto nos instrumentos de medição de vazão, como nos de pressão, além do fato de que a temperatura ambiente também variou durante o período de realização do ensaio. Conseqüentemente, foi necessário dividir o conjunto de dados em 3 subconjuntos distintos, de acordo com as fases de acompanhamento do processo, como foi explicado anteriormente no item 4.5.

Nota-se, na 1ª rodada de treinamento de rede neural apresentada na Tabela 4.8 e 4.9, que o 1º subconjunto já não identificou experimento suspeito. O 3º subconjunto, após a retirada do experimento 641 por ter sido identificado na 1ª rodada como suspeito, não apresentou mais nenhum experimento suspeito.

Entretanto, no 2º subconjunto, apresentado também na Tabela 4.8 e 4.9, foi mais difícil a eliminação dos experimentos suspeitos, necessitando de 5 rodadas para zerar o seu aparecimento. Na primeira rodada, foram identificados sete suspeitos, sendo que três dentre eles já haviam sido indicados na Tabela 4.3, ou seja, 60% dos experimentos identificados pela análise de "cluster" foram apontados pela rede neural. Os experimentos indicados pela análise estatística foram estes três mais um, ou seja, 26,7% dos experimentos apresentados na Tabela 4.4.



Voltando, agora, a unir todo o novo conjunto de dados, com os experimentos dos 3 subconjuntos que restaram após a retirada dos experimentos com erros grosseiros, obtém-se um conjunto constituído por 574 experimentos. Isto indica que foram retirados 14 experimentos suspeitos de conterem erros grosseiros. Dentre eles, quatro que foram identificados na primeira rodada, são justificados pela análise de “cluster” e pela análise estatística. Ainda na primeira rodada, com relação aos outros quatro experimentos detectados, verificou-se que eles sofreram influência de mudanças nas condições de operação durante o ensaio (variação de fluxo ou de pressão), resultando em realização de amostragem fora da condição de estabilização do processo. Os demais experimentos, ou seja, os identificados nas rodadas posteriores, podem ser resultado do próprio processo de ajuste do modelo da rede aos dados experimentais. A abordagem por rede neural mostra uma capacidade de identificação de duas classes diferentes de erros: os ditos grosseiros e os dos desvios dos instrumentos, ou seja, sistemáticos.

Por outro lado, existem resultados da análise estatística que não foram identificados pela rede neural. O conjunto de dados pode conter erros estocásticos, que não precisam necessariamente ser considerados como suspeitos de conterem erros grosseiros, apenas por não obedecerem ao critério do nível de confiança. É inerente ao processo a ocorrência de variância estatística.

#### **5.4 Verificação e validação da metodologia através de um segundo conjunto de dados**

A verificação e validação da metodologia utilizada para detecção de erros grosseiros foram realizadas com um conjunto de dados experimentais mais complexo e extenso, onde além das variáveis de processo descritas no primeiro conjunto de dados, também existiam variações externas com a utilização de diferentes ultracentrífugas montadas com diferentes arranjos físicos internos. A

finalidade é otimizar o desempenho separativo desse equipamento, sendo os dados similares aos dados utilizados por Migliavacca (1999) no Capítulo 7. No entanto, cabe salientar que o conjunto de dados aqui utilizado apresenta uma melhor qualidade, em termos de distribuição no espaço de domínio das variáveis analisadas, como na confiabilidade dos valores obtidos.

Este segundo conjunto é formado por dados adquiridos em dez diferentes ensaios realizados com diferentes máquinas com o objetivo de avaliar a influência de duas características de montagem (A e B), sobre o desempenho separativo. O planejamento desses ensaios e a identificação de cada um se encontram na Tabela 5.1. Estes ensaios fazem parte de um programa de avaliação mais abrangente que se encontra ainda em curso. Estes ensaios foram escolhidos de tal forma a completarem um projeto de experimentos menor com duas variáveis (característica A e B), em três níveis, com uma repetição. Em cada um desses ensaios foram investigadas duas variáveis de processo, que apresentam grande influência no desempenho separativo da máquina, e que aqui serão denominadas de *variável A* e *variável B*. Cada uma delas está dentro de um domínio preestabelecido com intervalos pequenos o suficiente para serem tratadas como variáveis contínuas, mas condizentes com a precisão dos instrumentos. Essas variáveis de processo definem então o que chamamos de ponto ou condição de operação.

TABELA 5. 1 :Distribuição e Identificação dos ensaios utilizados no segundo conjunto de dados

Montagem	Característica A	Característica B	Ensaio	
1	1	1	L2-08/99	L2-15/02
2	1	3	L1-02/00	L2-18/02
3	3	1	L2-09/99	L2-16/02
4	3	3	L1-04/00	L4-01/02
5	2	2	L2-10/99	L2-17/02

Em cada ponto de operação, definido pelas variáveis de processo, foram então coletados pelo menos três conjuntos de dados, juntamente com um par de amostras, para determinação das razões de abundância do produto  $R_p$  e

do rejeito  $R_w$  e cálculo dos parâmetros de separação, gerando ao todo 930 experimentos. Dos 930 experimentos realizados, 6 apresentaram condições insuficientes para realização de análise por espectrometria de massa da amostra. Portanto, também neste caso, esses experimentos foram excluídos imediatamente, resultando em um conjunto de 924 experimentos. Esse conjunto de 924 experimentos foi submetido à análise prévia através da verificação do balanço material macroscópico, quando foram excluídos 172 experimentos por inconsistência. Esse conjunto final de 752 experimentos foi então apresentado a uma rede neural segundo o esquema apresentado na Figura 5.1.

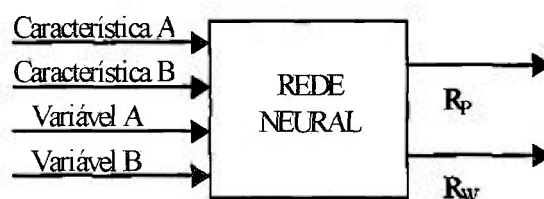


FIGURA 5. 1: Esquema das variáveis utilizadas na apresentação à rede neural

No presente caso, não foram realizadas análises prévias pelo método de "clusters", nem por métodos estatísticos. A proposta foi efetuar a detecção dos erros grosseiros pela rede neural, e somente depois, buscar justificativas para os experimentos apontados.

Foram testadas redes com 6, 8 e 10 neurônios na camada oculta com até 50000 apresentações para o treinamento da rede. A Tabela 5.2 apresenta os valores dos erros totais mínimos (dados pela somatória das diferenças entre os valores experimentais e os calculados pela rede de todo o conjunto de dados) encontrados quando permitidas até 50000 apresentações.

TABELA 5. 2 : Erro total mínimo encontrado para cada rede neural aplicada permitidas até 50000 apresentações no treinamento

No. de neurônios na camada interna	Erro total mínimo	No. de apresentações para o erro mínimo
6	1,5758	4300
8	1,1107	13500
10	1,1325	26200

Por meio desses valores, temos a indicação de que a rede que seria a melhor candidata à modelagem dos dados experimentais é a de 8 neurônios. O desenvolvimento dos erros totais ao longo do treinamento dessas redes é mostrado nas Figuras 5.2 e 5.3, onde se observa que, após 5000 apresentações, o erro decai lentamente com a continuidade do treinamento das redes. Para a identificação de erros grosseiros, esse refinamento no treinamento não é interessante.

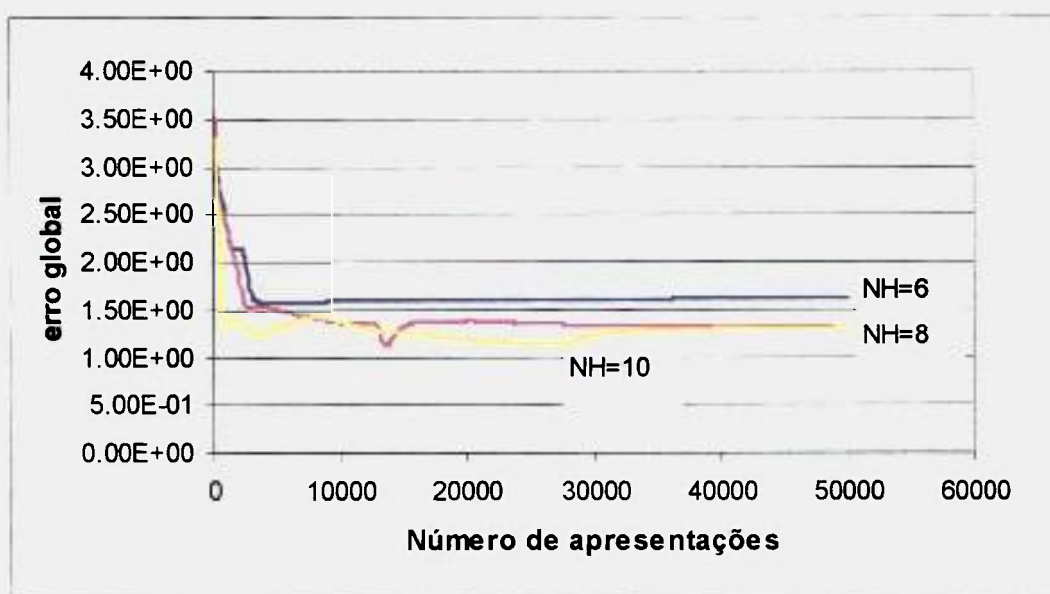


FIGURA 5. 2: Erro total das redes com 6, 8 e 10 neurônios ao longo de seus treinamentos até 50000 apresentações

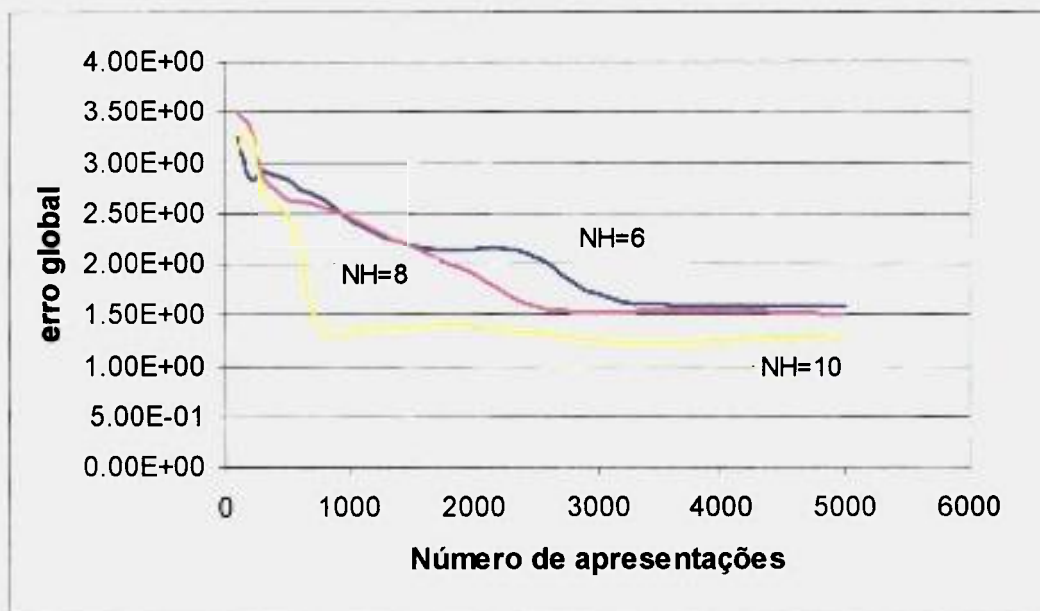


FIGURA 5. 3: Erro total das redes com 6, 8 e 10 neurônios ao longo de seus treinamentos até 5000 apresentações

Para cada rede avaliada, foi realizado o mesmo procedimento descrito na metodologia deste trabalho, através do tratamento dos resíduos entre os valores experimentais e os valores calculados pela rede neural das duas variáveis de saída definidas na Figura 5.1,  $R_P$  e  $R_W$ , onde então, foi aplicado o teste descrito por Himmelblau (1970). Assim o processo de detecção de erros foi avaliado nessas redes quando permitidas 5000, 3000 e um valor mínimo de apresentações, definido através da Figura 5.3 pelo final do forte decaimento inicial de cada rede. Nesse sentido, o valor mínimo de apresentações definido para cada rede é indicado juntamente com os demais resultados nas Tabelas 5.3, 5.4 e 5.5.

Os experimentos indicados nas Tabelas 5.3, 5.4 e 5.5 são quase todos identificados como erros grosseiros na análise de “cluster”, com exceção dos experimentos 54, 56 e 58, que foram apontados pela rede de 10 neurônios.

TABELA 5. 3 : Identificação de possíveis erros grosseiros através de diferentes redes neurais sendo o número de neurônios na camada oculta NH = 6  
Primeira rodada

NH	6		6	
No. de apresentações	3000		5000	
c	3.125170		3.125170	
Variável	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$
Mínimo	-0.0098	-0.0061	-0.0088	-0.0061
Máximo	0.0122	0.0118	0.0124	0.0110
Média	-0.0007	0.0023	-0.0004	0.0021
Sx	0.0032	0.0026	0.0032	0.0025
Média + Sx	0.0094	0.0103	0.0097	0.0100
Média - Sx	-0.0109	-0.0058	-0.0104	-0.0058
Experimentos suspeitos identificados	352	297	352	300
		299		301
		300		661
		301		
		302		
		597		
	661			

TABELA 5. 4 : Identificação de possíveis erros grosseiros através de diferentes redes neurais sendo o número de neurônios na camada oculta NH = 8  
Primeira rodada

NH	8		8		8	
No. de apresentações	2500		3000		5000	
c	3.125170		3.125170		3.125170	
Variável	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$
Mínimo	-0.0091	-0.0061	-0.0085	-0.0062	-0.0090	-0.0060
Máximo	0.0120	0.0117	0.0120	0.0110	0.0115	0.0102
Média	-0.0004	0.0021	-0.0003	0.0020	-0.0008	0.0018
Sx	0.0032	0.0025	0.0032	0.0025	0.0032	0.0025
Média + Sx	0.0097	0.0100	0.0097	0.0098	0.0093	0.0095
Média - Sx	-0.0105	-0.0058	-0.0103	-0.0059	-0.0109	-0.0058
Experimentos suspeitos identificados	57	297	57	297	57	300
	352	299	352	299	352	301
		300		300		661
		301		301		
		302		597		
		597		661		
	661					



TABELA 5. 5 : Identificação de possíveis erros grosseiros através de diferentes redes neurais sendo o número de neurônios na camada oculta NH = 10  
Primeira rodada

NH	10		10		10	
No. de apresentações	1000		3000		5000	
c	3.125170		3.125170		3.125170	
Variável	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$
Mínimo	-0.0087	-0.0048	-0.0087	-0.0056	-0.0090	-0.0067
Máximo	0.0120	0.0104	0.0112	0.0085	0.0108	0.0081
Média	-0.0003	0.0017	-0.0011	0.0009	-0.0012	0.0011
Sx	0.0032	0.0022	0.0033	0.0021	0.0033	0.0021
Média + Sx	0.0097	0.0087	0.0092	0.0075	0.0089	0.0078
Média - Sx	-0.0102	-0.0053	-0.0113	-0.0057	-0.0114	-0.0056
Experimentos suspeitos identificados	352	397	56	297	54	300
		299	57	299	56	661
		300	352	300	57	
		301		301	58	
		302		597	352	
		597				

Em seguida, os experimentos identificados nas tabelas acima, foram retirados do conjunto de dados experimentais e esses reapresentados à sua respectiva rede neural para nova rodada de busca de erros grosseiros, cujos resultados estão apresentados nas Tabelas 5.6, 5.7 e 5.8.

TABELA 5. 6 : Identificação de possíveis erros grosseiros através de diferentes redes neurais sendo o número de neurônios na camada oculta NH = 6  
Segunda rodada

NH	6		6	
No. de apresentações	3000		5000	
c	3.125043		3.125100	
Variável	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$
Mínimo	-0.0114	-0.0039	-0.0088	-0.0042
Máximo	0.0086	0.0096	0.0094	0.0099
Média	-0.0014	0.0024	-0.0004	0.0020
Sx	0.0033	0.0025	0.0032	0.0024
Média + Sx	0.0088	0.0102	0.0095	0.0096
Média - Sx	-0.0117	-0.0054	-0.0104	-0.0056
Experimentos suspeitos identificados	0	0	0	297
				299
				597

TABELA 5. 7 : Identificação de possíveis erros grosseiros através de diferentes redes neurais sendo o número de neurônios na camada oculta NH = 8  
Segunda rodada

NH	8		8		8	
No. de apresentações	2500		3000		5000	
c	3.125029		3.125043		3.125086	
Variável	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$
Mínimo	-0.0086	-0.0044	-0.0085	-0.0043	-0.0090	-0.0041
Máximo	0.0091	0.0079	0.0090	0.0096	0.0087	0.0090
Média	-0.0003	0.0017	-0.0004	0.0018	-0.0009	0.0017
Sx	0.0032	0.0023	0.0032	0.0024	0.0032	0.0024
Média + Sx	0.0096	0.0090	0.0095	0.0092	0.0091	0.0091
Média - Sx	-0.0102	-0.0056	-0.0103	-0.0056	-0.0108	-0.0057
Experimentos suspeitos identificados	0	0	0	302	0	0

Em alguns casos, nenhum suspeito foi apontado, encerrando o processo de busca. Quando novos suspeitos são apontados, o processo de exclusão e nova rodada foram repetidos até que não houvesse mais experimentos suspeitos apontados.

TABELA 5. 8 : Identificação de possíveis erros grosseiros através de diferentes redes neurais sendo o número de neurônios na camada oculta NH = 10  
Segunda rodada

NH	10		10		10	
No. de apresentações	1000		3000		5000	
c	3.125058		3.125043		3.125058	
Variável	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$	$e_{NRP}$	$e_{NRW}$
Mínimo	-0.0086	-0.0049	-0.0087	-0.0057	-0.0095	-0.0041
Máximo	0.0094	0.0070	0.0092	0.0072	0.0079	0.0076
Média	-0.0002	0.0015	-0.0012	0.0009	-0.0016	0.0012
Sx	0.0032	0.0021	0.0032	0.0020	0.0032	0.0021
Média + Sx	0.0097	0.0082	0.0089	0.0072	0.0082	0.0077
Média - Sx	-0.0101	-0.0051	-0.0113	-0.0055	-0.0115	-0.0054
Experimentos suspeitos identificados	0	0	54 58	302 661	0	0



Na Tabela 5.9 é mostrado o quadro resumo das identificações de erros grosseiros por experimento indicado. O número apresentado no quadro se refere à rodada em que o experimento foi apontado como suspeito. A ausência de números significa que o experimento não foi indicado como suspeito pela respectiva rede neural.

TABELA 5. 9 : Identificação de possíveis erros grosseiros através de diferentes redes neurais

NH	No. Apres.	Experimento apontado por $e_{NRP}$					Experimento apontado por $e_{NRW}$						
		54	56	57	58	352	297	299	300	301	302	597	661
6	3000					1	1	1	1	1	1	1	1
	5000					1	2	2	1	1		2	1
8	2500			1		1	1	1	1	1	1	1	1
	3000			1		1	1	1	1	1	2	1	1
	5000			1		1			1	1			1
10	1000					1	1	1	1	1	1	1	
	3000	2	1	1	2	1	1	1	1	1	2	1	2
	5000	1	1	1	1	1			1				1

A rede com 6 neurônios na camada oculta resultou praticamente nas mesmas indicações tanto quando permitidas 3000 apresentações, quanto com 5000 apresentações, sendo que com um menor número de apresentações a finalização foi imediata. O mesmo observamos na rede com 8 neurônios com treinamento de até 3000 apresentações. Essa rede com 5000 apresentações já sugere a acomodação da rede, ou seja o refinamento do treinamento, perdendo então a indicação de alguns suspeitos em  $e_{NRW}$ , que pela análise de "cluster" podem ser considerados erros grosseiros. Já a rede com 10 neurônios, mostrou resultados próximos às redes com 6 e 8 neurônios, somente quando treinada até o mínimo de apresentações, a saber, até 1000 apresentações. Com um número maior de apresentações essa rede passou a apontar experimentos em  $e_{NRP}$  que não se apresentam tão fortemente como suspeitos de conterem erros grosseiros como os anteriores. Com 5000 apresentações deixou de identificar importantes suspeitos, novamente sugerindo um superajuste para essa finalidade.

Os experimentos considerados críticos são 57, 352, 297, 299, 597 e 661. Esses experimentos são os mais fortemente identificados pela análise de

“cluster” como contendo erros grosseiros. Os experimentos 54, 56, 58, 300, 301 e 302 podem ser justificados pela análise de “cluster” porém, como se encontram nas proximidades dos limites do domínio das variáveis de entrada, a sua identificação é mais delicada. Assim, face ao discutido acima e considerando os valores dos erros globais observados nas rede ao final de um longo treinamento, podemos definir a rede com 8 neurônios na camada oculta com 2500 ou com 3000 apresentações como sendo a que melhor identificou os suspeitos de conterem erros grosseiros.

Esses resultados nos remetem à premissa original de escolha da rede para a tarefa de identificação de erros grosseiros, partindo de uma rede com um número de neurônios na camada interna mínimo para um bom ajuste dos dados experimentais após um treinamento completo, e refazendo o treinamento com um número mínimo de apresentações, de modo que o decaimento abrupto inicial do erro total seja superado. Com esse treinamento são obtidos os valores dos erros das variáveis analisadas em termos das diferenças entre os valores experimentais e os valores calculados e então aplicado um teste estatístico para a identificação de erros grosseiros, no presente caso um teste de hipótese para médias, descrito por Himmelblau (1970).

## 6 CONCLUSÕES

O tratamento convencional do conjunto de dados através da verificação do balanço material macroscópico, da análise por “clusters” e da análise estatística dos erros, nos indicou um grupo de experimentos considerados suspeitos de conterem erros grosseiros. O primeiro desses tratamentos é imperativo e deve anteceder qualquer outro tipo de análise, pois exclui os experimentos com inconsistência física de resultados. As análises, estatística e por “clusters”, foram realizadas de uma maneira muito criteriosa e, portanto, muito extensa.

No primeiro caso estudado, a detecção dos erros grosseiros através do uso de rede neural, conduziu a resultados pouco satisfatórios, comparados aos identificados pelas análises convencionais. Vários fatores podem ter influenciado neste resultado, como por exemplo a utilização de uma variável de processo não primária como o  $\delta U$ .

No segundo caso, os resultados se apresentaram bem mais consistentes, conduzindo a uma boa concordância entre os experimentos identificados pela análise de “cluster” e os detectados pela rede neural. Os “clusters” devem apresentar uma grande homogeneidade entre seus elementos, sendo assim, bastante sensíveis a uma discrepância que possa ocorrer entre os dados avaliados. Isto prova que a rede conseguiu identificar com boa exatidão, estes experimentos suspeitos de conterem erros grosseiros por não se ajustarem aos seus similares do mesmo grupo. Uma comparação direta entre a análise estatística empregada neste trabalho e a análise via rede neural não foi possível, desde que somente a rede neural foi capaz de identificar grupos com diferentes comportamentos dentro de um mesmo conjunto de dados.

Entretanto, com relação à análise estatística, tanto no primeiro como no segundo caso, não foi possível identificar todos os experimentos suspeitos por ela detectados. A presença de erros grosseiros no conjunto de dados deteriora a estimativa efetuada, ou seja, a própria média ou o desvio padrão, carregam a dispersão que os erros grosseiros possam causar. Em geral, a possibilidade de

falha na detecção de erros grosseiros através de métodos convencionais, aumenta com a quantidade desses erros presentes.

Para uma decisão final na eliminação dos dados suspeitos, deve ser considerada a experiência do pesquisador/ engenheiro, principalmente em casos como este, onde se busca um refinamento dos dados para posterior modelagem do comportamento separativo de um protótipo de ultracentrífuga, que é parte de projeto maior de pesquisa. Entretanto, a aplicação de rede neural mostrou ser uma excelente ferramenta para uma rápida detecção de erros grosseiros, pois requer pouco conhecimento fenomenológico do processo, executando a análise de dados de forma efetiva, e em menos tempo.

A identificação de erros grosseiros foi efetuada através da utilização de redes neurais, que se mostrou uma ferramenta muito ágil e versátil. A definição da estrutura da rede neural a ser empregada se baseia em algumas premissas:

- o número de neurônios da camada oculta deve respeitar a constatação de Pollard (1992), que diz que os resultados finais de uma rede neural de três camadas é pouco sensível ao número de neurônios após um mínimo; no entanto não deve possuir um número excessivo de neurônios para não provocar um superajuste da rede. Como regra, pode-se escolher a rede cujo erro global é mínimo após o treinamento completo para a definição de um modelo numérico;
- a quantidade de apresentações deve ser a menor possível após a acomodação inicial dos pesos, verificada através das curvas de acompanhamento do erro global do treinamento da rede. No presente caso, verificou-se que um número máximo de apresentações adequado para a detecção de erros foi da ordem de 3000. Treinamentos mais extensos causam a adaptação da rede aos experimentos com erros, o que pode acarretar um superajuste da rede;
- após o treinamento da rede, devem ser obtidos os desvios de cada uma das variáveis de resposta em termos da diferença entre os valores experimentais e os valores calculados pela rede; nesses conjuntos de desvios são então aplicados os testes estatísticos para identificação de

erros grosseiros, como o utilizado no presente trabalho, descrito por Himmelblau (1970);

O presente trabalho estabeleceu um procedimento para a indicação de dados suspeitos de conterem erros grosseiros dentro de um conjunto de dados onde as técnicas usuais para identificação de erros grosseiros não podem ser aplicadas ou sua utilização demanda um trabalho excessivamente longo. Além disso, também foi capaz de identificar erros sistemáticos provenientes dos instrumentos. Este trabalho possibilitou a depuração de conjuntos de dados, os quais estão sendo ou serão utilizados em novos trabalhos.

Como em toda e qualquer análise de dados experimentais, o julgamento do experimentador deve sempre ser cauteloso. Todas as indicações de possíveis erros grosseiros devem ser verificadas e justificadas dentro do conjunto global. O conhecimento do analista sobre o processo deve ser profundo e sua opinião pessoal deve ser levada em conta (Himmelblau, Cap.4.7, 1970). A decisão final sobre a exclusão ou não dos dados suspeitos de conterem erros grosseiros é do experimentador, que deve ponderar seus conhecimentos sobre o processo e as indicações efetuadas pela rede neural, ou por qualquer outra ferramenta.

Sugere-se como proposta para trabalhos futuros, o estudo de novos critérios de detecção de erros grosseiros em conjuntos de dados provenientes de ensaios de separação isotópica, considerando-se ainda, o uso de redes neurais, com outro protótipo de ultracentrífuga, e com maior redundância dos dados experimentais.

Seria interessante, também, explorar melhor o uso da rede neural como ferramenta de identificação de erros sistemáticos derivados de progressiva degradação da qualidade dos resultados do processo, culminando na aferição de instrumentos que fazem a medição das variáveis de processo.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Alves, R.B.M.; Nascimento C.A.O. Gross Errors Detection of Industrial Data by Neural Network and Cluster Techniques. *Brazilian J. of Chem. Eng.*, v.19, n.4, p.483-489, 2002.
- Avery, D.G.; Davis, E. *Uranium Enrichment by Gas Centrifuges*. London: Mills and Boon, 1973.
- Bhat, N.; McAvoy, T. Use of neural nets for dynamic modeling and control of chemical process systems, *Computers & Chemical Engineering*, v.14(4/5), p.573-583, 1990.
- Benedict, M.; Pigford, T.H.; Levi, H.W. *Nuclear Chemical Engineering*. 2.ed., New York, NY: McGraw-Hill, 1981.
- Bülau, H.C.; Ulrich, J.; Guardani, R.; Nascimento, C.A.O. Application of neural networks to data from a melt crystallization process for the detection of outliers. In: AIDA – International Seminar On Advances Data Analysis, 1999. Washington. **Proceedings...** Washington, 1999.
- Cohen, K. *The theory of isotope separation as applied to the large scale production of U<sup>235</sup>*. New York, McGraw-Hill, 1951.
- Himmelblau, D.M. *Process Analysis by Statistical Methods*. New York, NY: Wiley, 1970.
- Hoskins, J.C.; Himmelblau, D.M. Artificial neural network models of knowledge representation in chemical engineering, *Computers & Chemical Engineering*, v.12(9/19), p.881-890, 1988.
- Jordan, I. *Separação dos isótopos de urânio pelo processo da centrifugação em fase gasosa*. São Paulo, SP, 1966. Tese (Cátedra) – Escola Politécnica da Universidade de São Paulo.

- Kai, T. Theoretical research on gas-centrifugal separation for uranium enrichment. *J. Nucl. Sci. and Technol.*, 26(1): 157-160, 1989.
- Lapponi, J.C. Estatística usando o Excel. São Paulo, SP: Lapponi Treinamento e Editora, 2000.
- Migliavacca, S.C.P; Nascimento, C.A.O.; Rodrigues, C. Use of neural network for the simulation of a gas centrifuge. *J. Nucl. Sci. Technol.*, v.36, n.4, p.364-370, 1999a.
- Migliavacca, S.C.P. *Modelagem do comportamento separativo de ultracentrífugas via rede neural*. São Paulo, SP, 1999b. Tese (Doutorado) - Instituto de Pesquisas Energéticas e Nucleares.
- Migliavacca, S.C.P; Nascimento, C.A.O.; Rodrigues, C. Analyses and optimization of gas-centrifugal separation of uranium isotopes by neural networks. *J. Nucl. Sci. Technol.*, 2000.
- Nascimento, C.A.O. *NEURO14: neural networks modeling training program for multilayer feedforward networks using: backpropagation algorithm; momentum algorithm or; backpropagation with dynamic dumping factor algorithm*. Author: Claudio A. Oller do Nascimento. LSCP-DEQ/EPUSP (VERSION 14 on 29.9.1996 (Roberto Guardani).
- Nascimento, C.A.O., Giudici, R., Neural Network Based Approach for Optimization Applied to Industrial Nylon-6,6 Polymerization Process. *Comput. & Chem. Engng.*, No. 22, S595 (1998).
- Nascimento, C.A.O., Giudici, R. and Guardani, R., Neural Network Based Approach for Optimization of Industrial Chemical Processes. *Comput. & Chem. Engng.*, No. 24, 2303-2314 (2000).
- Olander, D.R., Technical Basis of the Gas Centrifuge. *Adv. Nucl. Sci. Technol.*, No. 6, 105-174 (1972).

- Plácido, J. *Desenvolvimento e aplicação industrial da técnica de reconciliação de dados*. São Paulo, SP, 1995. Dissertação (Mestrado) – Escola Politécnica da Universidade de São Paulo.
- Pollard, J.F.; Broussard, M.R.; Garrison, D.B.; San, K.Y. Process identification using neural networks, *Computers & Chemical Engineering*, v.16(4), p.253-270, 1992.
- Rumelhart, D. and McClelland, J., *Parallel Distributed Processing Explorations in the Microstructure of Cognition*, vol. 1, cap. 8, MIT, Cambridge, Mass. (1986).
- Soubbaramayer, Centrifugation. In: VILLANI, S., ed. *Uranium Enrichment*. Springer Verlag, Berlin, p. 183- 243 (1979).
- Zippe, G., *The Development of Short Bowl Ultracentrifuges*. Final report ORO-315 (1960).