

**UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA QUÍMICA
DESENVOLVIMENTO DE PROCESSOS QUÍMICOS**

**Aplicação das Técnicas de Redes Neurais
e de Análise de Componentes Principais na Modelagem
de uma Lagoa Aerada da RIPASA S/A**

Karla Patrícia Santos Oliveira

Autora

Milton Mori

Orientador

Roy Edward Bruns

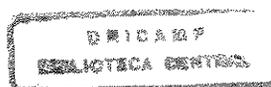
Co-Orientador

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

Campinas - São Paulo

Fevereiro/2000

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE



2000.7514

UNIDADE OC
CHAMADA:
T/UNICAMP
O14a
41551
278/00
PREÇO R\$ 11,00
DATA: 13.07.00
Nº CPD

CM-00143159-3

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

O14a Oliveira, Karla Patrícia Santos
Aplicação das técnicas de redes neurais e de análise de componentes principais na modelagem de uma lagoa aerada da RIPASA S/A / Karla Patrícia Santos Oliveira.--Campinas, SP: [s.n.], 2000.

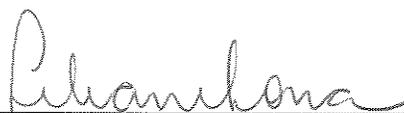
Orientadores: Milton Mori e Roy Edward Bruns.
Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Química.

1. Redes neurais. 2. Análise de componentes principais. 3. Papel - Indústria. 4. Meio ambiente. I. Mori, Milton. II. Bruns, Roy Edward. III. Universidade Estadual de Campinas. Faculdade de Engenharia Química. IV. Título.

Dissertação de Mestrado defendida por Karla Patrícia Santos Oliveira e aprovada em
28 de Fevereiro de 2000 pela banca examinadora constituída pelos doutores:



Prof. Dr. Milton Mori



Prof^a Dr^a Liliane Maria Ferrareso Lona Batista



Dr. Manoel Araújo Lourenço Telhada

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

**Este exemplar corresponde à redação final da Dissertação
de Mestrado em Engenharia Química**


Orientador

*Dedico este trabalho ao meu esposo - Vitaly,
aos meus pais – Roseana e Alencar –
aos meus irmãos – Erika, Line, Diego e Polly
e à minha avó – Benedita,
por inundarem minha vida de amor
e pelo apoio para realização de mais um sonho.*

AGRADECIMENTOS

À minha família que tanto me incentivou na escolha deste caminho.

Ao meu orientador pela amizade e pelo apoio e suporte não só no desenvolvimento deste trabalho como também nos demais desenvolvidos na área ambiental.

Ao meu co-orientador Roy Edward Bruns pela amizade e pelo apoio, incentivo e segurança transmitidos.

Aos alunos Cristian, Carlos e Luciana Ferrari por todo o incentivo prestado quanto à aprendizagem da técnica de Redes Neurais.

A Carlos Gontarski, Vitaly Félix Rodríguez Esquerre, Prof. Dr. Ronei Jesus Poppi, Prof^a. Dr^a. Liliane Lona e Prof. Dr. Edson Tomás pelas frutíferas discussões durante todo o trabalho.

Novamente ao meu esposo Vitaly Félix Rodríguez Esquerre pelo suporte nas horas de “desespero” quanto ao entendimento das técnicas e desenvolvimento do trabalho.

Aos professores e funcionários da Faculdade de Engenharia Química, em especial à Anna Ritta, Eunir, Rosa e Rita, por todo o auxílio prestado.

À RIPASA S/A – Indústria de Papel e Celulose, nas pessoas de Manoel Telhada e Jonas Vitti, pelo auxílio na pesquisa do sistema de lagoas aeradas de sua estação de tratamento de efluentes e pelo fornecimento de dados.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, pelo apoio financeiro.

Aos Profs. Selêude Wanderley da Nóbrega e Roberaldo Carvalho de Souza, da graduação, e à Sra. Jaguaraci Barros pelo incentivo para continuação dos meus estudos.

À turma do laboratório – LMSPQ - pelo apoio e companheirismo.

Aos meus antigos amigos pelo carinho transmitido mesmo à distância, e aos novos pela gentil e maravilhosa acolhida à nova faculdade e cidade.

Àqueles que, mesmo por trás dos “bastidores”, participaram da elaboração deste trabalho.

RESUMO

O uso industrial dos corpos hídricos, rios e lagos, tem sido motivo de grande preocupação nacional e mundial visto às condições de captação de água e de destinação da grande quantidade de resíduo gerada diariamente. A atividade industrial da RIPASA S/A não foge a esta regra, assim vê-se a necessidade de um trabalho constante de controle e otimização do seu sistema de tratamento de efluente.

Este trabalho teve como objetivo geral fornecer subsídios à RIPASA S/A de prever a demanda bioquímica de oxigênio (DBO₅) em situações futuras de despejo de sua lagoa aerada II.

Utilizou-se assim as técnicas de Redes Neurais Artificiais (RNAs) e de Análise de Componentes Principais (PCA) para modelagem do sistema e pré-processamento dos dados, respectivamente. Tomou-se um período de dois anos para a formação do banco de dados, sendo analisados inicialmente os seguintes parâmetros: DBO₅, demanda química de oxigênio (DQO), vazão, produção de papel e celulose, pH e sólidos suspensos. Os pacotes computacionais NeuroSolutions e Statistica foram utilizados.

Como o tempo de residência do sistema de lagoas aeradas da RIPASA S/A não era conhecido, trabalhou-se com as possibilidades de três a sete dias, tendo o de seis dias apresentado melhores resultados. A vasta quantidade de lacunas no banco de dados impossibilitou o uso dos parâmetros pH e sólidos suspensos reduzindo-se o número de dados disponíveis para a formação do modelo.

Os resultados da modelagem via RNA foram comparados com o obtido através de um dos métodos clássicos de modelagem, PCR (*Principal Component Regression*), e verificou-se que nenhuma das técnicas foi capaz de prever bem o comportamento do sistema em estudo quando utilizadas separadamente. Entretanto, uma boa performance da RNA foi obtida quando realizado um pré-processamento dos dados de entrada antes de serem alimentados à rede.

Desta forma a ortogonalização de variáveis via PCA é apresentada como uma importante ferramenta de auxílio na definição da topologia das RNAs e no seu mapeamento não linear.

ABSTRACT

In recent years, computer-based methods have been applied to many areas of environmental chemistry. In the process industry the use of modern control strategies is required due to increasing demands on the quality of its effluent treatment systems.

In this work a wastewater treatment system of a pulp and paper industry has been studied using Artificial Neural Networks (ANN) and the Principal Components Analysis to predict output environment parameters (BOD). Control process data sets generated from input and output of the current treatment system (an aerated lake) are used in this research. Variation within sampling of some auxiliary and process parameters including chemical oxygen demand (COD), biochemical oxygen demand (BOD), flow, pulp and paper production, pH and suspended solids are evaluated over a two-year period.

Predictive models are presented calculated from ANN and Principal Component Regressions (PCR) for the estimation of biochemical oxygen demand, one of the main process control variables.

The results show that neither principal component regression nor artificial neural network treatment is satisfactory when used separately in modeling and simulation. Neural network presents superior results for the training set but poorer ones than those from PCR for the test set. One explanation is that there are too few data resulting in an overfit of the training set.

Best prediction performance is achieved when the data are preprocessed using PCA, before they are fed to a backpropagated neural network composed of three neurons in a hidden layer and the Delta-Bar-Delta (DBD) learning algorithm. The PCA technique orthogonalizes the input original variables and helps the ANN nonlinear mapping.

SUMÁRIO

	p.
RESUMO	
ABSTRACT	
1 INTRODUÇÃO	1
1.1 OBJETIVOS ESPECÍFICOS	2
2 REVISÃO BIBLIOGRÁFICA	2
2.1 EFLUENTES LÍQUIDOS	3
2.1.1 Tratamento de efluentes	6
2.1.2 Lagoa Aerada	7
2.1.2.1 Parâmetros de monitoração	8
2.2 CASO EM ESTUDO: SISTEMA DE LAGOAS AERADAS DA RIPASA S/A	10
2.2.1 Estudos Iniciais – GONTARSKI et al. (1998)	10
2.3 REDES NEURAIS ARTIFICIAIS (RNAs)	11
2.3.1 Introdução	11
2.3.2 Um breve histórico	12
2.3.3 Redes neurais biológicas	13
2.3.4 Redes neurais artificiais	14
2.3.4.1 Processos de aprendizado	17
2.3.5 Mapeamento não linear multivariado	17
2.3.5.1 Analogia com ajuste polinomial	18
2.3.5.2 Funções erro e treinamento da rede	19
2.3.6 Perceptron multi-camadas	19
2.3.6.1 Arquitetura do perceptron multi-camadas	20
2.3.7 Algoritmos de treinamento	22
2.3.7.1 Algoritmo <i>backpropagation</i>	23
2.3.8 Redes neurais na engenharia química	27
2.3.8.1 Estudos Iniciais – GONTARSKI et al. (1998): Predição da DBO e DQO na saída da lagoa aerada II da RIPASA S/A.	28
2.4 ANÁLISE DE COMPONENTES PRINCIPAIS (PCA)	29
2.4.1 Introdução	29
2.4.2 Um breve histórico	29
2.4.3 Desenvolvimento Matemático	30
2.4.4 <i>Principal Component Regression</i> (PCR)	32
2.4.5 Utilização da técnica de PCA para pré-processamento dos <i>inputs</i> das RNAs	33
3. MATERIAIS E MÉTODOS	34
3.1. FERRAMENTAS COMPUTACIONAIS	34
3.2. METODOLOGIA	34
3.2.1 Pesquisa bibliográfica	34
3.2.1.1 A RIPASA S/A	35

3.2.2.2 Detecção de <i>outliers</i>	37
3.2.2.3 Determinação dos parâmetros a serem utilizados para a predição da DBO_{out}	38
3.2.3 Desenvolvimento do modelo	39
3.2.3.1 Número de amostras e parâmetros	39
3.2.3.2 Divisão de dados para treinamento, validação e teste	40
3.2.3.3 Determinação da topologia da rede	40
3.2.3.4 Treinamento e Teste da rede	42
3.2.3.5 Interpretação e integração do modelo	42
4 RESULTADOS E DISCUSSÕES	42
4.1 DETERMINAÇÃO DO TEMPO DE RESIDÊNCIA	42
4.2 RNA	44
4.3 PCA e RNA	48
4.4 PCR	55
5 CONCLUSÕES	57
6 SUGESTÕES	
7 REFERÊNCIAS BIBLIOGRÁFICAS	58
APÊNDICE	
1A: Gráfico comparativo entre DQO de entrada e DQO de saída para um período de cinco dias de residência	63
1B: Ajuste da rede para dados da DBO do conjunto de treino (tempo de residência de 5 dias).	64
1C: Ajuste da rede para dados da DBO do conjunto de teste (tempo de residência de 5 dias).	64

Lista de Figuras

	p.
Figura 2.1: O neurônio biológico.	14
Figura 2.2: Conexões sinápticas.	14
Figura 2.3: Modelo de McCulloch-Pitts.	15
Figura 2.4: Algumas funções de ativação: (a) linear, (b) thershold, (c) thershold linear e (d) sigmoidal.	16
Figura 2.5: Esquema de uma função não linear geral mapeada a partir de um conjunto de variáveis de input x_1, \dots, x_d e um conjunto de variáveis de output $\hat{Y}_1, \dots, \hat{Y}_c$.	18
Figura 2.6: Estrutura de uma rede neural simples.	20
Figura 2.7: Estrutura de uma rede neural contendo uma camada intermediária.	21
Figura 2.8: Desenho esquemático da superfície da função erro $E(w)$.	22
Figura 3.1: Sistema de tratamento de efluentes da RIPASA S.A.	36
Figura 4.1: Gráficos comparativos entre DQO de entrada e DQO de saída, e entre suas cargas, para os períodos de três a sete dias de residência.	43

Figura 4.1: Gráficos comparativos entre DQO de entrada e DQO de saída, e entre suas cargas, para os períodos de três a sete dias de residência.	43
Figura 4.2: Gráficos da variação de E do conjunto de teste da CDBO _{out} com o aumento do número de neurônios para os algoritmos Quickpropagation e Delta-Bar-Delta.	45
Figura 4.3: Ajuste da rede para os dados da CDBO _{out} do conjunto de treino e teste (tempo de residência de 6 e 7 dias).	47
Figura 4.4: Gráficos dos loadings para os tempos de residência de 6 dias e 7 dias.	50
Figura 4.5: CDBO _{out} medida e estimadas usando-se PCA-RNA dos conjuntos de teste correlacionados para os tempos de residência de 6 dias (a) e 7 dias (b).	51
Figure 4.6: Erro quadrático médio vs n° de componentes principais vs n° número de neurônios intermediários para os conjuntos de treino (a) e teste (b).	53
Figura 4.7: CDBO _{out} medida vs estimada usando-se PCA-RNA (arquivos de treino e teste – 6 dias).	54
Figura 4.8: Correlação entre CDBO _{out} estimada e medida pelos métodos RNA, PCA-RNA e PCR ² (conjunto de treino).	55
Figura 4.9: Correlação entre CDBO _{out} estimada e medida pelos métodos RNA, PCA-RNA e PCR ² (conjunto de teste).	56

Lista de Tabelas

Tabela 2.1: Principais contaminantes em tratamento de águas residuárias	p. 4
Tabela 2.2: Características físicas, químicas e biológicas das águas residuárias e suas fontes	5
Tabela 3.2: Grupos de variáveis utilizados na predição da DBO da saída da lagoa aerada II.	39
Tabela 4.1: Divisão de dados para composição dos arquivos de treino e teste.	44
Tabela 4.2: Melhores resultados na predição da CDBO _{out} , considerando-se tempos de residência de 3 à 7 dias.	46
Tabela 4.3: Melhores resultados obtidos para a modelagem.	48
Tabela 4.4: Análise da variância acumulada das componentes principais.	49
Tabela 4.5: Resultados para a predição da CDBO _{out} por PCA-RNA (algoritmo DBD).	52
Tabela 4.6: Resultados para a predição da CDBO _{out} utilizando as técnicas PCA-RNA (6dias).	52
Tabela 4.7: Resultados para a predição da CDBO _{out} .	56

1. INTRODUÇÃO

Juntamente com o aumento da atividade industrial, cresce também a demanda por meios mais eficientes de tratar os resíduos não aproveitáveis gerados durante os processos de produção. Os efluentes da indústria de papel e celulose são motivos de grande preocupação visto que podem conter produtos químicos nocivos ao meio ambiente que são introduzidos durante a produção (YANG, 1996). Esta preocupação tem se refletido na necessidade de se desenvolver sistemas de prevenção e controle para situações críticas.

Em termos legislativos, no Brasil, a DBO de 5 dias é um dos parâmetros utilizados para a determinação da matéria orgânica degradável de forma aeróbia de um efluente. Medidas para controle da qualidade do resíduo tratado são dificultadas tendo em vista este tempo assim, se fazem necessários estudos a partir de técnicas para predição imediata deste parâmetro.

A modelagem tradicionalmente usada em bioprocessos é baseada em equações de balanço aliadas às equações de taxa de crescimento microbiano, de consumo de substrato ou formação de produtos, entre outras (HAMMER, 1979). Um sistema de tratamento de efluentes é demasiadamente complexo e, além disso, está sujeito a perturbações externas, como o clima. Desta forma, a falta de informação sobre o mecanismo do processo pode dificultar ainda mais a utilização da modelagem determinística tradicional (COTE et al., 1995).

Neste contexto, as técnicas de Redes Neurais Artificiais (RNAs) têm apresentado bons resultados como ferramentas de modelagem para sistemas cujos mecanismos internos não são suficientemente conhecidos (MORRIS et al., 1994), como é o caso do tratamento de resíduos (GONTARSKI et al., 1998; GONTARSKI et al., 1999; HÁCK E KOHNE, 1996; PU et al., 1995). Elas possuem a propriedade de - uma vez apresentadas a um conjunto de dados de entrada e saída - “aprender” as relações que estivessem implícitas entre os dados contidos nesse conjunto.

Quanto maior o número de variáveis, maior é o número de parâmetros a serem ajustados pela rede e, conseqüentemente, maior o seu tempo de convergência. Assim várias técnicas de compressão de dados têm sido pesquisadas. No caso de modelagem de sistemas ambientais a técnica estatística de Análise de Componentes Principais (PCA) tem sido

estudada em grande escala (JACKSON, 1991; PIELOU, 1984). Esta apresenta duas vantagens principais no pré-processamento de dados: redução da dimensionalidade e ortogonalidade.

Através do uso de dados ortogonais como entrada da rede é possível otimizar o seu tamanho e conseqüentemente diminuir o número de parâmetros a serem ajustados (KANJILAL, 1995). Para métodos estatísticos convencionais de regressão, como Multiple Linear Regression – MLR, o uso de variáveis ortogonais também é desejável visto que reduz o erro de modelagem.

Os méritos do uso de RNAs em combinação com a técnica estatística de PCA têm sido apresentados em vários trabalhos, tais como: CANCELLA e FANG (1996), DUTTA et al. (1997), HOLCOM e MORARI (1992), KOMPANY-ZARED et al. (1998), PONTON e KLEMES (1993), ZUPAN e GASTEIGER (1993), etc.

1.1 OBJETIVOS ESPECÍFICOS

Esta pesquisa tem como objetivo principal o uso das técnicas de PCA e RNAs na modelagem do sistema de lagoas aeradas da RIPASA S/A para a predição da DBO no efluente tratado. Para tanto, foram investigados os principais parâmetros de controle deste sistema, a saber: demanda química e bioquímica de oxigênio, vazão do efluente, pH, sólidos suspensos, produção de papel e de celulose. O objetivo secundário foi verificar a relação entre os parâmetros investigados, incluindo a DBO.

2. REVISÃO BIBLIOGRÁFICA

Neste capítulo apresenta-se a revisão bibliográfica relacionada ao sistema em estudo a ser modelado, sendo detalhadas as técnicas utilizadas para a realização desta modelagem.

Inicia-se com efluentes líquidos, apresentando-se suas características, formas de geração e conseqüências de seu lançamento em corpos receptores, verificando assim, a necessidade de seu tratamento. É apresentado sucintamente o processo anaeróbio de tratamento, dando maior ênfase ao aeróbio que é o utilizado pela RIPASA S/A, mais especificamente seu sistema de lagoas aeradas.

Posteriormente são relatadas as técnicas de PCA e RNAs, utilizadas no pré-processamento e modelagem do sistema de lagoas aeradas, respectivamente. Incluem-se suas formulações matemáticas, campos de aplicação, vantagens e desvantagens de uso.

2.1 EFLUENTES LÍQUIDOS

De acordo com METCALF & EDDY (1991), um efluente líquido pode ser definido como uma combinação de resíduos líquidos ou diluídos em água, removidos de residências, instituições, estabelecimentos comerciais e industriais, junto a águas subterrâneas, superficiais ou pluviais que possam estar presentes.

Muitos produtos orgânicos são descartados diariamente, provenientes de esgotos domésticos, efluentes industriais, atividades agropecuárias, produtos farmacêuticos, descartes de laboratórios, curtumes, refinarias de petróleo, etc. (CONNELL & MILLER, 1984).

O lançamento de compostos orgânicos em corpos receptores pode causar sérias conseqüências, comprometendo sua qualidade. MONTEIRO (1975) detalha o crescimento de bactérias oxidativas sob o aumento da concentração de matéria orgânica biodegradável, com conseqüente aumento do consumo de oxigênio.

Em condições normais, um rio é capaz de receber uma carga apreciável, eliminando-a gradativamente mediante ações naturais que se processam ao longo de vários quilômetros do seu percurso, entretanto um curso d'água desprovido de oxigênio dissolvido ocasiona a destruição dos organismos aeróbios e praticamente impossibilita o uso de suas águas para múltiplos usos e finalidades.

As águas doces, salobras e salinas do território brasileiro são protegidas pela legislação brasileira, portaria 36/GM de 18 de junho de 1986 (CONSELHO NACIONAL DO MEIO AMBIENTE, 1990). No âmbito estadual, SÃO PAULO (1986) estabelece padrões para efluentes de qualquer fonte poluidora quanto a pH, temperatura, materiais sedimentáveis, DBO₅ e concentrações de substâncias tóxicas, etc.

Para atender às exigências legislativas, as fontes poluidoras devem dispor de sistemas de tratamento de efluentes líquidos, e a forma de tratamento estará relacionada intimamente com as características dos efluentes. METCALF & EDDY (1991) citam os principais

contaminantes em tratamento de águas residuárias (Tabela 2.1) e as principais características físicas, químicas e biológicas das águas residuárias, bem como suas fontes (Tabela 2.2).

Tabela 2.1: Principais contaminantes em tratamento de águas residuárias.

CONTAMINANTE	IMPORTÂNCIA
Sólidos suspensos	Sólidos suspensos podem levar ao desenvolvimento de depósitos de lodo e condições anaeróbias quando o efluente líquido não tratado é lançado no ambiente aquático.
Orgânicos biodegradáveis	Compostos principalmente por proteínas, carboidratos e gorduras, os orgânicos biodegradáveis são quantificados basicamente em termos de DBO (demanda bioquímica de oxigênio) e DQO (demanda química de oxigênio). Se lançado sem tratamento ao ambiente, sua estabilização biológica pode levar à queda da reserva de oxigênio natural e ao desenvolvimento de condições sépticas.
Patogênicos	Algumas doenças podem ser transmitidas por organismos patogênicos em águas residuárias.
Nutrientes	Tanto nitrogênio quanto fósforo, juntos ao carbono, são nutrientes essenciais para o crescimento. Quando lançados no ambiente aquático, estes nutrientes podem levar ao crescimento de uma vida aquática não desejável. Quando lançados em excessivas quantidades sobre a terra, também podem poluir águas subterrâneas.
Poluentes perigosos	Compostos orgânicos e inorgânicos selecionados com base no conhecimento de apresentarem carcinogenicidade, mutagenicidade, teratogenicidade ou toxicidade. Muitos destes compostos são encontrados em águas residuárias.
Orgânicos refratários	Estes orgânicos tendem a resistir a métodos convencionais de tratamento de efluentes líquidos. Exemplos típicos incluem surfactantes, fenóis e pesticidas agrícolas.
Metais pesados	Metais pesados são geralmente adicionados às águas residuárias de atividades comercial e industrial e devem ser removidos se o efluente for reutilizado.
Inorgânicos dissolvidos	Constituintes inorgânicos como cálcio, sódio e sulfato são adicionados à água de abastecimento doméstico e devem ser removidos se o efluente for reutilizado.

Fonte: METCALF & EDDY (1991)

Tabela 2.2: Características físicas, químicas e biológicas das águas residuárias e suas fontes.

CARACTERÍSTICAS	FONTE
PROPRIEDADES FÍSICAS	
• Cor	Resíduos domésticos e industriais, degradação natural de materiais orgânicos
• Odor	Águas residuárias em decomposição e resíduos industriais
• Sólidos	Abastecimento de água potável, resíduos domésticos e industriais, erosão de solos, infiltração
• Temperatura	Resíduos domésticos e industriais
CONSTITUINTES QUÍMICOS	
• Orgânicos	
<i>Carboidratos</i>	Resíduos domésticos, comerciais e industriais
<i>Gorduras, óleos e graxas</i>	Resíduos domésticos, comerciais e industriais
<i>Pesticidas</i>	Resíduos agrícolas
<i>Fenóis</i>	Resíduos industriais
<i>Proteínas</i>	Resíduos domésticos, comerciais e industriais
<i>Poluentes perigosos</i>	Resíduos domésticos, comerciais e industriais
<i>Surfactantes</i>	Resíduos domésticos, comerciais e industriais
<i>Compostos orgânicos voláteis</i>	Resíduos domésticos, comerciais e industriais
<i>Outros</i>	Degradação natural de matérias orgânicos
• Inorgânicos	
<i>Alcalinidade</i>	Resíduos domésticos, abastecimento de água potável, infiltração de água subterrânea
<i>Cloretos</i>	Resíduos domésticos, abastecimento de água potável, infiltração de água subterrânea
<i>Metais pesados</i>	Resíduos industriais
<i>Nitrogênio</i>	Resíduos domésticos e agrícolas
<i>pH</i>	Resíduos domésticos, comerciais e industriais
<i>Fósforo</i>	Resíduos domésticos, comerciais e industriais, drenagem natural de água
<i>Poluentes perigosos</i>	Resíduos domésticos, comerciais e industriais
<i>Enxofre</i>	Abastecimento de água potável, resíduos domésticos, comerciais e industriais
• Gases	
<i>Gás sulfídrico</i>	Decomposição de resíduos domésticos
<i>Metano</i>	Decomposição de resíduos domésticos
<i>Oxigênio</i>	Abastecimento de água potável, infiltração com a superfície da água
CONSTITUINTES BIOLÓGICOS	
• Animais	Cursos d'água abertos e plantas de tratamento de efluentes
• Plantas	Cursos d'água abertos e plantas de tratamento de efluentes
• Protistas	
<i>Eubactéria</i>	Resíduos domésticos, infiltração com a superfície da água e plantas de tratamento de efluentes
<i>Arqueobactérias</i>	Resíduos domésticos infiltração com a superfície da água e plantas de tratamento de efluentes
<i>Vírus</i>	Resíduos domésticos

Fonte: METCALF & EDDY (1991)

Na Tabela 2.1 pode-se observar que a atividade industrial apresenta-se como uma considerável fonte potencial de poluição dos corpos hídricos. Os efluentes gerados na produção de papel e celulose constituem um dos maiores problemas, haja visto que poucas indústrias necessitam de tamanha quantidade de água para seu processo de manufatura como esta (GALVÃO, 1988). Estes efluentes são gerados diariamente e podem conter cargas químicas provenientes dos seus processos de produção (YANG, 1996).

Basicamente todo o resíduo proveniente da fabricação do papel e celulose é originado nos seguintes processos:

- máquina de papel;
- branqueamento da celulose, e
- fábrica de recuperação (YANG, 1996).

Este resíduo é composto por produtos químicos usados durante o processo como compostos organo-clorados, peróxidos, hipoclorito, sendo portanto bastante ácido (pH aproximadamente 3) e são produzidos em grande quantidade.

Os métodos adotados para tratamento destes efluentes são baseados em sistemas convencionais, isto é, tratamento físico, físico-químico e biológico, embora envolvam o manuseio de substâncias complexas contidas nos efluentes (GALVÃO, 1988).

2.1.1 Tratamento de efluentes

Para preservar a qualidade dos cursos d'água, a Legislação Federal Brasileira divide-os em classes de acordo com suas características. As legislações estaduais devem impor restrições aos parâmetros dos efluentes lançados a estes cursos d'água de forma a preservar sua classificação.

Um efluente que não atenda às condições que lhe são impostas deverá ser alterado, isto é, tratado. É necessária, então, a implantação de um sistema de tratamento adequado às características desse efluente. Os efluentes inorgânicos são tratados preferencialmente por processos físico-químicos. Os orgânicos podem receber tratamento por processo físico-químico ou biológico. Os processos biológicos são adequados quando o efluente orgânico apresenta características biodegradáveis.

Os processos biológicos que ocorrem na presença de oxigênio livre, denominados aeróbios, são os mais utilizados e conhecidos. A implantação de sistemas com processos aeróbios oferece, portanto, poucos riscos de investimento, pois têm custo baixo de construção e apresentam operação e manutenção fáceis e econômicas. Porém, são sistemas que ocupam grandes áreas, os equipamentos necessários como aeradores demandam energia e o lodo, gerado como subproduto, ainda é passível de decomposição por ação biológica.

Outro processo de decomposição biológica ocorre na ausência de oxigênio livre e denomina-se processo anaeróbio. Quando o tratamento é realizado em lagoas anaeróbias pode-se ter problemas estéticos devido a formação de espuma da superfície da água; já quando realizado em bio-reatores fechados, o processo pode apresentar as vantagens de produção do gás metano como subproduto que é utilizado como fonte de energia, menor área de ocupação, e ainda produção de lodo em menor quantidade e já estabilizado.

2.1.2 Lagoa aerada

As lagoas aeradas têm sido empregadas no tratamento de águas residuárias de indústrias de papel e celulose, de laticínios, de frigoríficos, de fábricas de conservas, entre outras.

Estas lagoas efetuam a depuração das águas residuárias através da remoção da matéria orgânica e de algumas substâncias inorgânicas pelo metabolismo de organismos diversos, em presença de oxigênio dissolvido na água. Elas possuem profundidade variando de 1,80 m a 4,5 m (dependendo do sistema de aeração) e trabalham com oxigênio mínimo dissolvido de 1,5 mg/L, este é introduzido na massa líquida através de processos mecânicos.

Os tempos de retenção, ou residência, variam de 2 a 10 dias, não sendo comum o desenvolvimento de algas. Existe também uma produção de lodo, o qual é retirado em decantadores, pois devido a agitação contínua na lagoa ele fica suspenso na água.

Em termos de redução de DBO, a eficiência das lagoas aeradas pode ser superior a 95%. Dentre os principais fatores que influenciam na atividade da lagoa têm-se:

- a) Fatores incontroláveis:
- temperatura;
 - agitação (ventos);
 - evaporação, e
 - chuvas.
- b) Fatores parcialmente controláveis:
- permeabilidade do fundo;
 - nutrientes, e
 - carga poluidora.
- c) Fatores relacionados ao projeto de construção da lagoa
- localização;
 - número de unidades;
 - disposição das unidades;
 - dimensões;
 - período de detenção, e
 - detalhes construtivos.

Estes fatores exercem influência determinante na eficiência do processo. Um estudo consistente destes fatores deve ser realizado assim como diversos parâmetros devem ser monitorados e controlados de modo a garantir o bom funcionamento das lagoas.

2.1.2.1 Parâmetros de monitoração

Para atender às condições requeridas pelos microrganismos é importante o monitoramento de alguns parâmetros. A escolha dos parâmetros de monitoração deve ser feita baseada no conhecimento do processo aeróbio. Eles devem ser suficientes para informar sobre a realidade do comportamento da biomassa e como esta reage ante variações das características da água residuária sob tratamento. Por fim, devem ser de fácil operacionalidade, uma vez satisfeita a exigência anterior.

Os principais parâmetros de monitoração são apresentados a seguir:

Demanda química de oxigênio

É determinada pela oxidação dos resíduos orgânicos em uma solução ácida de dicromato de potássio em ebulição. Este processo oxida quase todos os compostos orgânicos existentes na amostra analisada, liberando o gás carbônico e a água, cuja reação tem, geralmente, uma eficácia de mais de 95% de oxidação de todos os compostos existentes na água.

Este resultado é obtido rapidamente (3 horas), porém não especifica a proporção de matéria orgânica existente na água residual que pode ser oxidada por bactérias e nem a velocidade de bio-oxidação.

Demanda bioquímica de oxigênio

É a quantidade de oxigênio necessária para a oxidação da matéria orgânica pelas bactérias, isto é, determina-se a concentração de matéria orgânica existente na água residual que pode ser oxidada (biodegradada).

Como a DBO é uma análise que demora 5 dias, ela é geralmente estimada através de relações entre DBO e DQO para obter resultados aproximados mais rápidos. Daí vê-se a necessidade de um modelo para a determinação imediata deste parâmetro.

Pela legislação é necessário reduzir, pelo tratamento de efluentes, a carga orgânica de 60 a 80%.

pH

O pH é um fator importante no crescimento das bactérias, sendo que a maioria delas não tolera pH acima de 9,5 ou abaixo de 4,0. METCALF & EDDY (1991) citam a faixa ótima de pH para o crescimento de bactérias entre 6,5 e 7,5.

Sólidos suspensos

A variação da quantidade de sólidos suspensos está diretamente relacionada com a concentração da DBO na lagoa sendo este o parâmetro que exige um monitoramento mais árduo.

Vazão

Como o tempo de detenção da lagoa é praticamente constante, variando apenas com o decorrer do tempo pelo assoreamento da mesma, a sua capacidade de tratamento depende da vazão do efluente a ser tratado e das características deste.

Produção de papel e celulose

O monitoramento da produção de papel e celulose se faz necessária para controle tanto do processo de produção como do sistema de tratamento de efluentes haja visto que praticamente todo efluente industrial é gerado por estes dois processos.

2.2 CASO EM ESTUDO: SISTEMA DE LAGOAS AERADAS DA RIPASA S/A

Foi estudado o sistema de lagoas aeradas da RIPASA S/A logo, são apresentados alguns aspectos que motivaram os trabalhos iniciais para a modelagem deste.

2.2.1 Estudos Iniciais – GONTARSKI et al. (1998)

Como citado anteriormente, a DBO é um dos principais parâmetros utilizados para a determinação da matéria orgânica degradável de forma aeróbia de um efluente. Medidas para controle da qualidade do resíduo tratado via DBO são dificultadas tendo em vista o tempo de 5 dias, exigidos por lei, para a sua medição.

Um trabalho realizado pelo Laboratório de Modelagem e Simulação de Processos Químicos/UNICAMP juntamente com a RIPASA S/A surgiu em agosto de 1997 (GONTARSKI et al., 1998) na busca de ferramentas para que o operador pudesse agir mais eficientemente no sistema, basicamente através da adição de oxigênio e nutrientes, de forma que a variável principal a ser controlada (a DBO na saída para o rio) fosse mantida dentro dos limites estabelecidos por lei.

No trabalho supracitado é realizada a predição da DBO na saída do sistema de lagoas aeradas, por ser este responsável pela redução da DBO. A modelagem tradicional usada em bioprocessos, através de equações de balanço juntamente com equações de taxas específicas de crescimento microbiano, consumo de substrato ou formação de produtos, foi substituída pela modelagem via redes neurais artificiais tendo em vista a complexidade do sistema.

Uma descrição sucinta da metodologia, resultados e conclusões de GONTARSKI et al. (1998) são apresentados após a descrição detalhada das técnicas computacionais de modelagem utilizadas - redes neurais artificiais.

2.3 REDES NEURAIAS ARTIFICIAIS (RNAs)

2.3.1 Introdução

Redes neurais artificiais (RNAs) é o nome dado a um conjunto de métodos matemáticos e algoritmos computacionais especialmente projetados para simular o processamento de informações e aquisição de conhecimento do cérebro humano.

A partir dos anos 80, iniciou-se um crescimento acelerado nos níveis das atividades de pesquisa em relação às RNAs. Enquanto muitos pesquisadores têm se esforçado de maneira a desenvolver princípios fundamentais e novos algoritmos, os estudos também têm se voltado para o campo de aplicações práticas. De fato, nos últimos anos tem ficado claro que as redes neurais oferecem um conjunto poderoso de ferramentas para a solução de problemas de reconhecimento de padrões, processamento de dados, modelagem e controle não linear, que pode ser considerado como complementar aos métodos mais convencionais.

A seguir será apresentado um breve histórico e uma descrição detalhada dos usos e aplicações das técnicas de RNAs e o desenvolvimento matemático que as envolvem.

2.3.2 Um breve histórico

O cérebro humano é a estrutura computacional mais complexa conhecida pelo homem. Sua capacidade de aprender através de estímulos externos tem inspirado muitos pesquisadores das mais diversas áreas na tentativa de criar um modelo computacional que copie as atividades do cérebro de diferentes maneiras.

As origens das redes neurais artificiais data dos anos 40 com a publicação do trabalho de McCulloch e Pitts, *A Logical Calculus of Ideas Immanent in Nervous Activity*, que descrevia um modelo de cálculo lógico de uma rede neural.

O segundo grande passo ocorreu em 1949 com a publicação do livro *The Organization of Behaviour* por Hebb, em que ele propôs um mecanismo específico para aprendizagem em redes neurais, baseado nas conexões sinápticas biológicas.

No final da década de 50 Rosenblatt publicou o trabalho: *Principles of Neurodynamics*, que apresentou a estrutura de um perceptron, uma unidade de reconhecimento de padrões que é a base do neurônio de uma rede neural.

Por volta de 1969 Minsky and Papert publicaram o livro *Perceptrons* que mostrava as limitações de um perceptron isolado. Tal trabalho, adicionado ao alto custo computacional da época desestimulou pesquisadores e agências de financiamento a investirem em redes neurais, fazendo com que na década de 70 apenas aplicações de reconhecida eficácia, tais como reconhecimento de imagens, fossem desenvolvidas.

Na década de 80 as pesquisas em RNAs ressurgiram devido principalmente à publicação de importantes trabalhos tais como o da energia de aproximação de Hopfield, *Neural Networks and Physical Systems with Emergent Collective Computational Abilities*, e o algoritmo de aprendizagem de retropropagação (*back-propagation*) para perceptrons de múltiplas camadas primeiramente proposto por Werbos, modificado várias vezes, e então popularizado por Rumelhart e McClelland após a publicação do livro de dois volumes *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, em 1986.

Apesar do desenvolvimento de redes neurais ter sido retomado, na década de 80 os trabalhos de modelagem e simulação estavam voltados praticamente para a utilização da técnica de colocação ortogonal devido à sua característica de necessitar um número reduzido de pontos de discretização e exigir menor carga computacional de resolução numérica dos modelos matemáticos.

A década de 90 tem sido caracterizada pela consolidação dos fundamentos teóricos das RNAs assim como pelo seu sucesso em aplicações práticas, ocasionado pelo aumento do poder de cálculo dos microcomputadores e estações de trabalhos ('workstation') e redução do custo dos mesmos. Tenta-se, então, compensar o período de estagnação nesta área de pesquisa.

2.3.3 Redes neurais biológicas

Como o próprio nome indica, as redes neurais artificiais são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes, numa tentativa de imitar o funcionamento do cérebro humano. São formadas por unidades básicas que buscam copiar a forma de ligação dos neurônios no cérebro humano.

Para um melhor entendimento da base estrutural de uma RNA faz-se necessário uma breve abordagem de citologia.

O sistema nervoso é formado por um conjunto extremamente complexo de células, os neurônios. O neurônio, unidade celular fundamental do cérebro, é composto por três regiões distintas, soma, dentritos e axônio, que são apresentadas na Figura 2.1. A parte central da célula, ou soma, é a região nuclear onde está a maior parte do material intracelular e onde ocorrem os processos metabólicos. Do soma partem dois tipos de filamentos, os dentritos e o axônio. O neurônio recebe sinais (impulsos nervosos) de outros neurônios através dos dentritos (receptores) e transmite o sinal gerado através do seu corpo celular (transmissor). Vê-se então que os dentritos formam um conjunto de terminais de entrada e o axônio um longo terminal de saída.

A formação da estrutura nervosa ocorre através das conexões sinápticas, que conectam a árvore dendrítica com os axônios de outras células (Figura 2.2). Quando os impulsos nervosos chegam ao terminal sináptico, certas substâncias chamadas neurotransmissores alteram o potencial elétrico das membranas de modo a promover ou inibir a passagem do impulso nervoso. A efetividade de uma sinapse pode ser ajustada pela passagem de sinais através dela então, a sinapse pode aprender a partir das atividades em que ela participa.

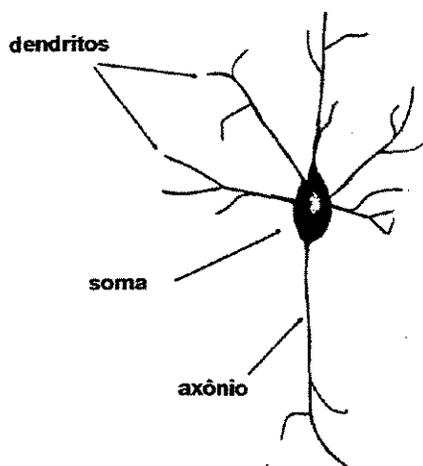


Figura 2.1: O neurônio biológico.

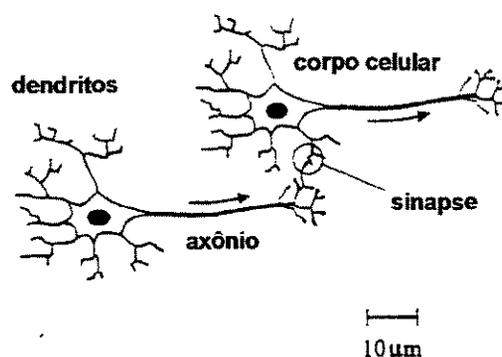


Figura 2.2: Conexões sinápticas.

O cérebro humano possui cerca de 10^{11} neurônios e estima-se que o número de conexões sinápticas gira em torno de 10^{14} a 10^{15} , possibilitando a formação de uma rede muito complexa (BISHOP, 1994). Pode-se dizer que o cérebro representa um sistema dinâmico de retro-alimentação, não linear, massivamente paralelo de proporções astronômicas e que, portanto, consegue realizar tarefas com uma velocidade surpreendentemente superior ao melhor computador.

2.3.4 Redes neurais artificiais

A RNA é um instrumento computacional que tenta imitar o sistema de aprendizagem do cérebro humano. A propriedade chave dos sistemas neurais biológicos e artificiais está na sua habilidade de modificar respostas como um resultado de exposição a

sinais externos. Isto é geralmente citado como aprendizagem e ocorre através da mudança da força das sinápses.

O neurônio artificial, proposto por MacCulloch e Pitts (1943), apresentado na Figura 2.3 pode ser considerado como uma função não linear que transforma um conjunto de variáveis de *input* x_i , ($i = 1, \dots, d$) em um conjunto de variáveis \hat{y} . Para distinguir um neurônio biológico do artificial, algumas vezes este último será chamado de unidade de processamento.

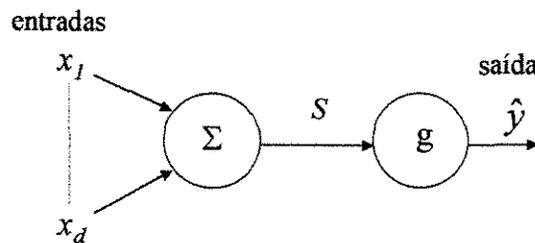


Figura 2.3: Modelo de McCulloch-Pitts.

No modelo de McCulloch-Pitts o sinal x_i do *input* i é primeiramente multiplicado por um parâmetro w_i conhecido como peso (que é análogo à intensidade sináptica do neurônio biológico) e é então somado a todos os *input* pesados para dar o *input* total para a unidade na forma

$$S = \sum_{i=1}^d w_i x_i + w_0, \quad (1)$$

onde o parâmetro de deslocamento w_0 é chamado de *bias* (e corresponde à descarga limite em um neurônio biológico). Formalmente, o *bias* pode ser considerado como um caso especial de um peso de um *input* extra cujo valor x_0 é permanentemente mantido em +1. Assim, pode-se escrever a Equação (1) na forma

$$S = \sum_{i=0}^d w_i x_i, \quad (2)$$

onde $x_0 = 1$. Como os pesos e os *bias* podem apresentar qualquer sinal, as sinápses podem ser excitadoras ou inibidoras. O *output* \hat{y} da unidade é então dado pela operação sobre uma função de ativação g , assim

$$\hat{y} = g(S). \quad (3)$$

Algumas funções de ativação são apresentadas na Figura 2.4. O modelo original de McCulloch-Pitts usou a função degrau mostrada na Figura 2.4 (b). Na prática, a função sigmoidal tem sido a função de ativação mais utilizada.

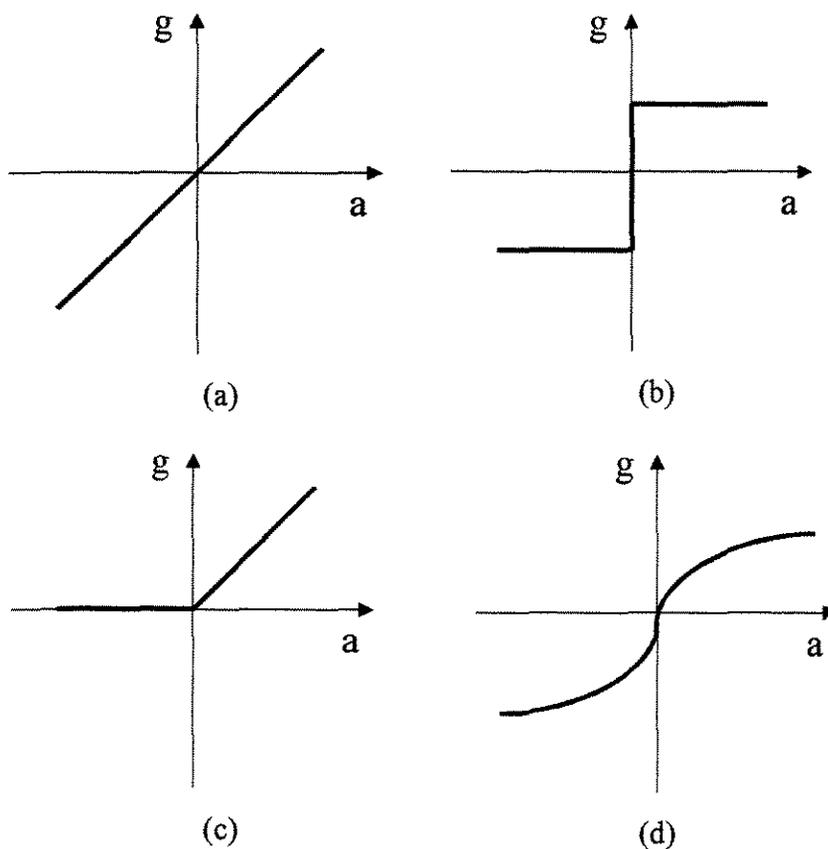


Figura 2.4: Algumas funções de ativação: (a) linear, (b) thershold, (c) thershold linear e (d) sigmoidal.

Nota-se então que este modelo simples de neurônios forma o elemento matemático básico dos modelos de redes neurais artificiais. Assim, através da conexão de vários elementos processadores é possível construir uma classe muito geral de mapeamento não linear, que pode ser aplicada aos mais variados problemas práticos. A adaptação dos valores dos pesos, de acordo com um apropriado algoritmo de treinamento, pode permitir a aprendizagem em resposta aos dados externos.

2.3.4.1 Processos de aprendizado

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

Denomina-se algoritmo de aprendizado um conjunto de regras bem definidas para a solução de um problema de aprendizado. Existem muitos tipos de algoritmos de aprendizado específicos para determinados modelos de redes neurais, estes algoritmos diferem entre si principalmente pelo modo que seus pesos são modificados.

Outro fator importante é a maneira pela qual uma rede neural se relaciona com o ambiente. Nesse contexto existem os seguintes paradigmas de aprendizado:

- Aprendizado supervisionado, quando é utilizado um agente externo que indica à rede a resposta desejada para o padrão de entrada;
- Aprendizado não supervisionado (auto-organização), quando não existe um agente externo indicando a resposta desejada para os padrões de entrada, e
- Reforço, quando um crítico externo avalia a resposta fornecida pela rede.

2.3.5 Mapeamento não linear multivariado

Como citado anteriormente, as RNAs formam a base para um grande número de aplicações práticas. A Figura 2.5 apresenta um esquema de uma função não linear que possui d variáveis independentes x_1, \dots, x_d e c variáveis dependentes $\hat{Y}_1, \dots, \hat{Y}_c$. De acordo com a nomenclatura utilizadas pelas redes neurais, os x 's são chamados de variáveis de

input e os \hat{Y}_k 's de variáveis de *output*. Cada \hat{Y}_k pode ser uma função não linear arbitrária dos *inputs*.

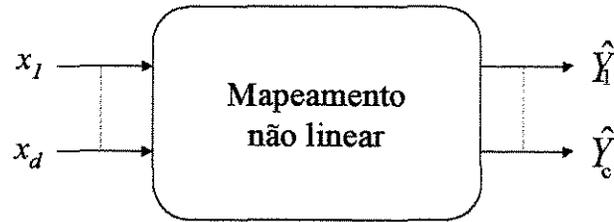


Figura 2.5: Esquema de uma função não linear geral mapeada a partir de um conjunto de variáveis de input x_1, \dots, x_d e um conjunto de variáveis de output $\hat{Y}_1, \dots, \hat{Y}_c$.

2.3.5.1 Analogia com ajuste polinomial

Fazendo-se uma analogia entre o treinamento de uma RNA e o problema de ajuste de uma curva usando um simples polinômio, considera-se por exemplo que este apresenta-se de ordem m dado por,

$$\hat{Y} = w_m x^m + \dots + w_1 x + w_0 = \sum_{j=0}^m w_j x^j. \quad (4)$$

A Equação (4) pode ser considerada como um mapeamento não linear que possui os x 's como variáveis de *input* para produzir o \hat{Y} como variável de *output*. A forma precisa da função $\hat{Y}(x)$ é determinada pelos valores dos parâmetros w_0, \dots, w_m que são análogos aos pesos na rede neural, sendo w_0 análogo ao parâmetro *bias* da Equação (1). Assim, este polinômio pode ser escrito na forma $\hat{Y} = \hat{Y}(x; w)$ como é feito para a maioria dos mapeamentos não lineares.

Existem dois fatores importantes em que as redes neurais diferem de um simples polinômio. Primeiramente, a RNA pode ter um número muito maior de variáveis de input e de output se comparado com o mapeamento polinomial. Em segundo lugar, a rede neural pode aproximar mais eficientemente uma classe muito grande de funções. Assim, para satisfazer as mesmas propriedades das RNAs, o mapeamento polinomial multivariado

necessita de um número de coeficientes ajustáveis exponencialmente maior que os parâmetros a serem ajustados pela rede (BISHOP, 1994).

2.3.5.2 Funções erro e treinamento da rede

O problema de determinação dos valores dos pesos de uma rede neural é chamado de treinamento. Uma função erro apropriada é definida com relação a um conjunto de dados, e a busca dos parâmetros (pesos) é feita de modo a minimizar esta função.

Cada vetor de *input* $x^q = (x_1^q, \dots, x_d^q)$ do conjunto de dados tem um *output* desejado y^q . O erro para a saída k quando a rede é apresentada com o padrão q é dada por $\hat{y}_k(x^q; w) - y_k^q$. O erro total para o conjunto padrão total pode então ser definido como o quadrado dos erros individuais somados sobre todos os *outputs* e padrões. Isto fornece uma função erro, para o uso no treinamento da rede neural, na forma

$$E = \frac{1}{2} \sum_{q=1}^n \sum_{k=1}^c [\hat{y}_k(x^q; w) - y_k^q]^2. \quad (5)$$

Embora sendo a soma quadrática dos erros a forma mais comumente usada de função erro, pode-se notar que em alguns casos é mais apropriado tomar a média individual do erro quadrático, ou seja, dividir o erro pelo número total de observações (E/n).

2.3.6 Perceptron multi-camadas (MLP)

Quando RNAs de uma só camada são utilizadas, o mapeamento é realizado diretamente entre as variáveis de *input* e as de *output*, ou seja, não é possível a formação de uma representação interna. Tal restrição implica que variáveis de *input* resultem em variáveis de *output* similares, o que leva o sistema à incapacidade de aprender importantes mapeamentos.

O desenvolvimento do algoritmo de treinamento *backpropagation* mostrou ser possível treinar eficientemente redes com camadas intermediárias, resultando assim no

modelo de RNAs mais utilizado atualmente, as redes Perceptron Multi-Camadas (MLP) (BISHOP, 1994).

Nessas redes, cada camada tem uma função específica. A camada de *output* recebe os estímulos da camada intermediária e constrói a variável que será a resposta. As camadas intermediárias funcionam como extratoras de características, seus pesos são uma codificação das características apresentados pelas variáveis de *input* e permitem que a rede crie sua própria representação.

2.3.6.1 Arquitetura do perceptron multi-camadas

Na seção 2.3.4 introduziu-se o conceito de uma única unidade de processamento, descrita pelas Equações (2) e (3). Considerando-se então um conjunto de m *outputs*, todos contendo *inputs* comuns, então é obtido uma rede neural tendo uma única camada de parâmetros adaptativos (pesos) como ilustrado na Figura 2.6. As variáveis de *output*, denotadas por \hat{y}_j , são dadas por

$$\hat{y}_j = g\left(\sum_{i=0}^d w_{ji} x_i\right), \quad (6)$$

onde w_{ij} é o peso do *input* i para a unidade j , e $g(\)$ é a função de ativação discutida anteriormente. Novamente pode-se incluir parâmetros *bias* como casos especiais de pesos provenientes de *inputs* extras $x_0 = 1$.

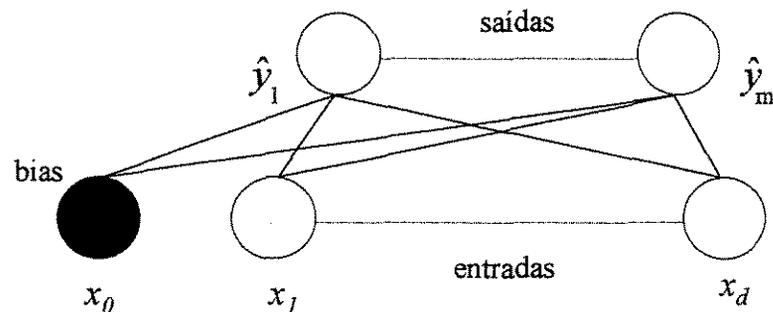


Figura 2.6: Estrutura de uma rede neural simples.

Na Figura 2.6, x_i representa os *inputs*, \hat{y}_j os *outputs* e as linhas que conectam as unidades de *input* e *output* representam os pesos correspondentes. O *input* extra $x_0 = 1$ é apresentado pelo círculo pintado e representa o parâmetro *bias*. As linhas que conectam o *bias* ao *output* representam os pesos, w_{j0} .

Enquanto o tipo de rede apresentado na Figura 2.6 tem algumas aplicações práticas, um classe muito mais poderosa de redes neurais é obtida se forem consideradas redes com pelo menos uma camada intermediária entre os *inputs* e *outputs*, como mostrado na Figura 2.7.

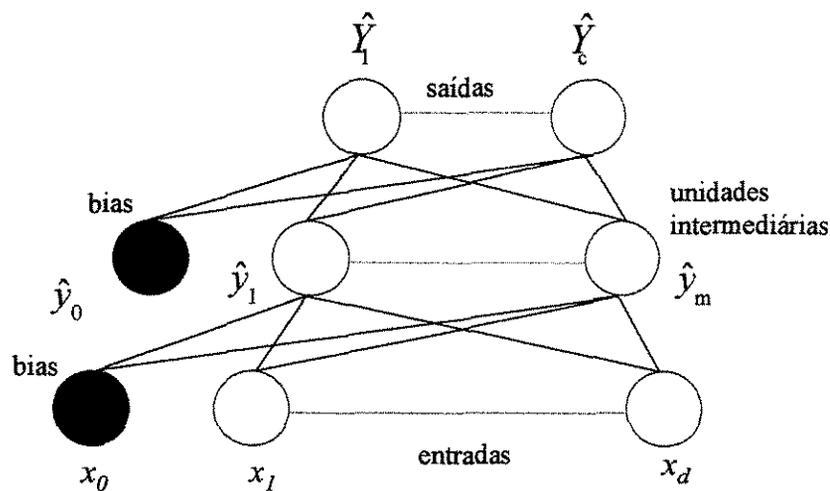


Figura 2.7: Estrutura de uma rede neural contendo uma camada intermediária.

A ativação das unidades da camada intermediária é dada pela Equação (2) e os *outputs* da rede são obtidos pela ativação sobre as saídas da camada intermediária, ou seja

$$\hat{Y}_k = G\left(\sum_{j=0}^m W_{kj} \hat{y}_j\right), \quad (7)$$

onde, W_{kj} representa os pesos na segunda camada conectando a unidade intermediária j à unidade de output k . Neste caso também pode ser introduzido uma unidade intermediária extra com ativação $y_0 = 1$ como o *bias* para as unidades de saída. O termo *bias* (para as unidades intermediárias e de *output*) têm o importante papel de garantir que as redes neurais

possam representar mapeamentos não lineares. Combinando-se então as Equações (6) e (7), tem-se a expressão completa para a transformação representada pela rede,

$$\hat{Y}_k = G \left(\sum_{j=0}^m W_{jk} g \left(\sum_{i=0}^d w_{ji} x_i \right) \right). \quad (8)$$

Novamente, cada componente da equação acima corresponde a um elemento do diagrama na Figura 2.7. Note que as funções de ativação g e G não precisam necessariamente ser a mesma.

2.3.7 Algoritmos de treinamento

A função erro pode ser considerada geometricamente como uma superfície ajustando-se sobre o espaço dos pesos, como mostrado na Figura 2.8, e o problema de treinamento da rede corresponde à pesquisa pelo mínimo global w_4 da superfície do erro.

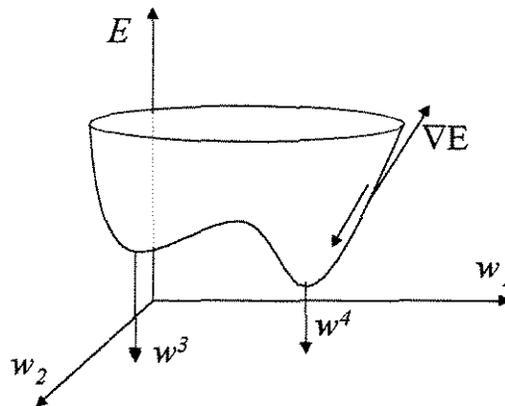


Figura 2.8: Desenho esquemático da superfície da função erro $E(w)$.

Para uma rede multi-camadas a função erro é altamente não linear em função dos pesos, e a pesquisa pelo mínimo geralmente procede de modo iterativo, começando por alguns pontos encontrados no espaço dos pesos. Alguns algoritmos encontram os mínimos locais mais próximos, enquanto outros são capazes de escapar dos mínimos locais e oferecer a possibilidade de encontrar o mínimo global. Em geral, a superfície de erro é extremamente

complexa e para muitas aplicações práticas um bom mínimo local pode ser suficiente para encontrar resultados satisfatórios.

Muitos algoritmos executam a minimização da função erro através das derivadas desta com relação aos pesos da rede. Estas derivadas formam as componentes do vetor gradiente $\nabla E(w)$ da função erro, que, em um dado ponto no espaço dos pesos, dá o gradiente da superfície do erro, como indicado na Figura 2.8.

Um dos mais importantes fatores da classe das funções de mapeamento não lineares dado pela MLP é que existe um procedimento computacional eficiente para avaliar as derivadas da função erro, baseadas na técnica de retropropagação do erro, ou seja, no algoritmo de *backpropagation*.

2.3.7.1 Algoritmo *backpropagation*

Durante o treinamento com o algoritmo *backpropagation*, a rede opera em uma sequência de dois passos. Primeiro, um conjunto de variáveis é apresentado à camada de *input* da rede. A atividade flui através da rede, camada por camada, até que a resposta seja produzida pela camada de *output*. No segundo passo, o *output* obtido é comparado ao *output* desejado para esse conjunto particular. Se esta não estiver correta, o erro é calculado. O erro é propagado a partir da camada de *output* até a camada de *input*, e os pesos das conexões das unidades das camadas internas vão sendo modificados conforme o erro retro-propagado.

Assim, vê-se que a formulação matemática do algoritmo de *backpropagation* consiste num problema de atualização dos pesos em função do erro.

Para a formulação matemática deste problema é considerada uma rede contendo apenas uma camada intermediária. Primeiramente tem-se que a função da soma quadrática dos erros - ou resíduos, Equação (5) - pode ser escrita como a soma dos quadrados dos resíduos em relação a cada padrão separadamente,

$$E = \sum_{q=1}^n E^q \quad E^q = \frac{1}{2} \sum_{k=1}^c [\hat{Y}_k(x^q; w) - y_k^q]^2, \quad (9)$$

onde $\hat{Y}(x;w)$ é dado pelo mapeamento da rede, Equação (8). Assim, para a obtenção da derivada total considera-se a derivada para cada padrão separadamente.

Nas próximas etapas do desenvolvimento matemático omiti-se o índice q referente aos pontos de dados. Considerando inicialmente as derivadas com relação aos pesos na segunda camada (a camada de pesos das unidades intermediárias para às de *output*), rescreve-se as variáveis de *output* da rede na forma,

$$\hat{Y}_k = G(A_k), \quad A_k = \sum_{j=0}^m W_{kj} \hat{y}_j. \quad (10)$$

As derivadas com relação aos pesos da camada de *output* podem então ser escritas como,

$$\frac{\partial E^q}{\partial W_{kj}} = \frac{\partial E^q}{\partial A_k} \frac{\partial A_k}{\partial W_{kj}}. \quad (11)$$

Introduzindo-se a definição,

$$\delta_k \equiv \frac{\partial E^q}{\partial A_k}. \quad (12)$$

Então, usando-se a Equação (10), tem-se a derivada na forma,

$$\frac{\partial E^q}{\partial W_{kj}} = \delta_k \hat{y}_j. \quad (13)$$

A expressão para δ_k pode ser encontrada pelas Equações (9), (10) e (12), logo

$$\delta_k = G'(A_k) [\hat{Y}_k - y_k]. \quad (14)$$

Como δ_k é proporcional à diferença entre o *output* da rede e o valor desejado, algumas vezes ele é citado como um erro. Note que, para as funções de ativação sigmoideal a derivada $G'(A)$ é facilmente expressada em termos de $G(A)$, ou seja

$$G'(A) = G(A) [1 - G(A)], \quad (15)$$

logo, o tempo computacional para gravação na implementação do algoritmo é pequeno.

De maneira a encontrar a expressão correspondente para a derivada com relação aos pesos na primeira camada, toma-se as ativações das unidades intermediárias na forma

$$\hat{y}_j = g(a_j), \quad a_j = \sum_{i=0}^d w_{ji} x_i. \quad (16)$$

Então a derivada requerida é escrita como,

$$\frac{\partial E^q}{\partial w_{ji}} = \frac{\partial E^q}{\partial a_j} \frac{\partial a_j}{\partial w_{ji}}. \quad (17)$$

Da Equação (16) nota-se que $\frac{\partial a_j}{\partial w_{ji}} = x_i$. Então definindo-se

$$\delta_j \equiv \frac{\partial E^q}{\partial a_j}, \quad (18)$$

a derivada tem a forma,

$$\frac{\partial E^q}{\partial w_{ji}} = \delta_j x_i. \quad (19)$$

Note que a equação acima tem a mesma forma que a derivada para os pesos da segunda camada dada pela Equação (13); assim, a derivada do erro para um dado peso que conecta uma unidade de *input* à uma unidade intermediária é dada pelo produto dos δ 's para a unidade intermediária e o valor da variável de *input*.

Finalmente, a expressão para os δ 's é obtida facilmente pela regra de derivadas parciais,

$$\delta_j = \frac{\partial E^q}{\partial a_j} = \sum_{k=1}^c \frac{\partial E^q}{\partial A_k} \frac{\partial A_k}{\partial a_j}. \quad (20)$$

Usando-se as Equações (10), (12) e (16) obtém-se

$$\delta_j = g'(a_j) \sum_{k=1}^c W_{kj} \delta_k. \quad (21)$$

A expressão da Equação (21) pode ser interpretada em termos do diagrama da rede como uma propagação de sinais do erro, dado por δ_k , retro-propagando através da rede partindo dos pesos da camada de *output*, sendo esta a origem do termo *backpropagation*.

Um importante fator para os cálculos das derivadas é sua eficiência computacional.

Geralmente o número de pesos é muito maior que o número de unidades de processamento, e a contribuição dominante do erro propagado ou retro-propagado é proveniente da avaliação da soma dos pesos (com a avaliação da função de ativação sendo negligenciada). Entretanto, a técnica de *backpropagation* permite que todas as derivadas sejam avaliadas usando uma única direção de propagação, ou seja, a retro-propagação, através do uso das Equações (13) e (19).

2.3.8 Redes neurais na engenharia química

A grande maioria dos processos químicos apresenta fortes não linearidades que dificultam bastante a obtenção de modelos fenomenológicos, a ponto de em alguns casos conseguir-se apenas um modelo extremamente simplificado de validade em condições específicas.

É exatamente neste nicho que as redes neurais surgem como uma técnica interessante de ser empregada devido à sua característica de aprender o que ocorre no processo. Neste sentido diferencia-se das técnicas convencionais de ajuste não linear de uma curva a um conjunto de pontos experimentais. Desta forma pode-se contornar, em muitos casos, as dificuldades de obtenção de modelos fenomenológicos representativos do processo.

Entretanto, o sucesso da modelagem e controle de processos via rede neural depende fortemente do conhecimento das principais variáveis do processo, além de se ter uma base de dados de boa qualidade, ou seja, que contenha todas as informações importantes do processo e o domínio desejado.

Embora existam muitos tipos de redes neurais, duas classes de rede são geralmente empregadas em engenharia química: a rede de múltiplas camadas (MLP), cuja informação segue uma única direção (*feedforward*), e a rede recorrente. Redes *feedforward* possuem a característica de modelarem bem processos estacionários enquanto que as redes recorrentes são melhores para modelar processos dinâmicos. Na modelagem de processos biológicos a rede *feedforward* tem apresentado grande aplicação (GONTARSKI et al., 1998; GONTARSKI et al., 1999; KARIM e RIVERA, 1992; PSICHOGIOS e UNGAR, 1992).

As etapas para desenvolvimento de um modelo de rede podem ser descritas como: análise do processo e construção da base de dados, configuração de uma estrutura de rede neural, treinamento ou aprendizado da rede e, por último, a validação do modelo no sentido de conseguir prever novas situações dentro do domínio de dados experimentais estudados.

2.3.8.1 Estudos Iniciais – GONTARSKI et al. (1998): Predição da DBO e DQO na saída da lagoa aerada II da RIPASA S/A.

As técnicas de redes neurais artificiais foram utilizadas na predição da DBO e DQO na saída de um sistema de lagoas de aeradas da RIPASA S/A. Foram estudadas variações das amostras de demanda bioquímica de oxigênio (DBO), demanda química de oxigênio (DQO), vazão e produção, obtidas durante o período de janeiro de 1993 à abril de 1997.

Devido ao assoreamento da lagoa aerada I, ocasionado pelo lodo formado durante o tratamento, e portanto mudança das suas dimensões, um estudo inicial teve que ser realizado de maneira a obter-se a confirmação do tempo de residência do efluente no sistema de lagoas. Essa informação é fundamental para esse tipo de estudo, pois, como esse é um sistema dinâmico onde, os valores dos parâmetros de entrada coletados em um determinado dia correspondem aos de saída coletados alguns dias depois.

Sabendo-se que o tempo de residência médio de um efluente em uma lagoa aerada encontra-se entre 2 e 10 dias, os dados foram correlacionados para períodos de 3, 4, 5, 6 e 7 dias, na tentativa de sua determinação. A melhor sobreposição dos gráficos tipo “tempo *versus* valores de entrada e saída da DQO” foi obtido para um tempo de 5 dias.

Com o tempo de residência estimado em 5 dias o arquivo de teste foi montado utilizando-se 20% dos dados (34), escolhidos aleatoriamente, sendo o restante (141) como arquivo de treinamento.

As variáveis de entrada (*inputs*) da rede foram DBO e DQO de entrada da lagoa I, vazão de entrada e saída e produção, e as de saída (*outputs*) DBO, DQO de saída da lagoa II. Seguindo-se a literatura (COTE et al., 1995; HÄECK e KOHONE, 1996; PU e HUNG, 1995) foi trabalhado com uma camada intermediária apenas.

A determinação do número de neurônios da camada interna e do algoritmo foi realizada de forma conjunta. Estudou-se redes para os algoritmos *QuickPropagation (QP)*, *Delta-Bar-Delta (DBD)* e *Online Backpropagation Randomized (OLBR)*. Para cada um dos algoritmos, o número de neurônios foi variado de 1 até 10 e foi feito um treinamento com 100000 iterações, totalizando-se 30 estruturas de redes estudadas.

Testando o conjunto de dados de treinamento e teste respectivo ao período de cinco dias para redes de diferentes topologias, determinou-se que a melhor estrutura para treinamento correspondeu a 3 neurônios na camada intermediária e o algoritmo de

retropropagação *On Line Backpropagation Randomized*, a partir da comparação entre os erros quadráticos médios obtidos para o conjunto de teste. Contudo, o erro ainda foi muito alto (0,156), significando que apesar da rede poder ajustar o conjunto de treinamento, ela ainda é incapaz de prever com confiabilidade o comportamento de um novo conjunto de dados originários da mesma fonte que os de treinamento.

Tal resultado foi justificado pela descontinuidade dos dados fornecidos pela empresa, que já formavam um conjunto limitado, sendo reduzido ainda mais na definição do tempo de residência.

Em continuidade ao compromisso firmado entre a RIPASA S/A e o Laboratório de Modelagem e Simulação de Processos Químicos / UNICAMP, estudos em busca de novas técnicas que pudessem auxiliar na modelagem do sistema de lagoas supracitado foram realizados. Uma técnica estatística de ortogonalização de variáveis, como a análise de componentes principais, a ser utilizada conjuntamente com as RNAs pareceu ser a melhor alternativa tendo em vista os trabalhos de CANCELLA e FANG, 1996; HOLCOMB e MORARI, 1992; KANJILAL, 1995; KOMPANY-ZAREH et al., 1998, entre outros.

2.4 ANÁLISE DE COMPONENTES PRINCIPAIS (PCA)

2.4.1 Introdução

A análise de componentes principais PCA (do inglês *Principal Component Analysis*), também chamada de transformação canônica (DUTTA et al., 1998), apresenta como idéia básica encontrar combinações lineares das variáveis de um sistema, de modo a reduzir sua dimensão a uma coleção muito menor de variáveis transformadas, que ainda descreva suas características.

2.4.2 Um breve histórico

Os procedimentos gerais da técnica de componentes principais, na forma que é conhecida hoje, foram desenvolvidos por Harold Hotelling e publicado no seu trabalho pioneiro em 1933, embora esta tenha sido originada em 1901 por Karl Pearson (MARDIA et al., 1979; JACKSON, 1991).

O desenvolvimento desta técnica foi um tanto fragmentada nos anos seguintes, alternando-se entre períodos de estagnação e de grande atividade. Com o advento computacional surgiu um grande número de estudos voltados para aplicação da técnica de PCA nas mais diversas áreas por apresentarem a necessidade de compreensão e interpretação de dados de seus sistemas multivariados.

Para um melhor entendimento da técnica de PCA, na seção à seguir apresenta-se uma descrição clara do seu desenvolvimento matemático.

2.4.3 Desenvolvimento Matemático

PCA é uma técnica estatística para redução da dimensionalidade de um conjunto de dados multivariados.

Considerando-se uma matriz \tilde{X} , $N \times P$, onde N é o número de amostras e P o número de variáveis independentes, as componentes principais são obtidas diagonalizando a matriz de covariância $\tilde{X}^t \tilde{X}$, onde \tilde{X}^t é a transposta de \tilde{X} .

Os autovetores correspondentes da matriz de covariância formam os eixos das coordenadas do sistema do espaço transformado e as variâncias nos dados ao longo destes eixos são os autovalores, λ_i . ADAMS (1995) e JACKSON (1991) apresentam formas simples para cálculo dos autovetores e autovalores.

Uma transformação linear simples relaciona os valores das componentes principais \tilde{Y} , chamadas *scores*, aos valores dos dados originais, \tilde{X} ,

$$\tilde{Y} = \tilde{X} \tilde{W}. \quad (22)$$

Geometricamente as componentes principais (eixos do espaço transformado) são relacionadas com os eixos do espaço original pelas rotações dadas pelos *loadings* ou pesos, \tilde{W} . Os *loadings* e os *scores* podem ser normalizados de modo que a soma quadrática de cada um seja igual a seus apropriados autovalores e a unidade, respectivamente, ou vice e versa.

A transformação do conjunto de dados pelas componentes principais apresenta duas vantagens principais para o tratamento de dados, a redução da dimensionalidade e a ortogonalidade das variáveis.

A primeira componente principal (CP) é orientada na direção que contém a maior variância dos dados, veja MARDIA et al. (1979). A segunda CP é ortogonal à primeira e é orientada na direção que contém a máxima quantidade de variância residual dos dados, isto é, a variância não explicada pela primeira CP. Sucessivamente, as demais componentes explicam a quantidade máxima de variância residual.

Quanto maior a porção de variância nos dados puder ser explicada pelas primeiras componentes principais, uma redução de dimensionalidade mais eficiente pode ser obtida. Esta porção da variância é representada pela razão da soma dos autovalores das componentes no espaço reduzido, $(i=1, 2, \dots, k)$ e da soma total de autovalores $(j=1, 2, \dots, p)$, ou seja,

$$V_k = \left(\frac{\lambda_1 + \dots + \lambda_k}{\lambda_1 + \dots + \lambda_p} \right), \quad (23)$$

onde k e p representam as primeiras k e todas as componentes principais, respectivamente.

Nem sempre a escolha das primeiras componentes principais levam a melhor representação de um sistema, haja vista que os últimos podem conter informações importantes de algumas variáveis originais (DESPAGNE e MASSART, 1998; MARDIA et al., 1979). Assim os autores têm sugerido diversas formas para seleção das componentes principais, dentre elas tem-se:

- (a) incluir apenas CPs suficientes para explicar 90% do total da variância. E se a matriz de correlação foi utilizada, excluir aquelas CPs cujos autovalores são menores que um (MARDIA et al., 1979);
- (b) utilizar critérios estatísticos (GONZÁLEZ e GONZÁLEZ-ARJONA, 1995; JACKSON, 1991), e
- (c) analisar gráficos dos autovalores, *SCREE plot*, proposto por Cattell (1966) (JACKSON, 1991; MARDIA et al., 1979).

Como citado anteriormente, além da redução da dimensionalidade a técnica de PCA apresenta a vantagem de fornecer um novo conjunto de parâmetros não correlacionados.

A propriedade de ortogonalidade das CPs é desejável de maneira a reduzir o erro da modelagem em tratamentos estatísticos, especialmente quando as CPs apresentam-se física, química ou biologicamente facilmente interpretáveis.

2.4.4 Principal Component Regression (PCR)

As RNAs não são as únicas técnicas para modelagem multivariada não linear. Alguns métodos estatísticos clássicos encontram seu uso na modelagem de processos, sendo os mais utilizados: Principal Component Regression - PCR e Partial Least Squares regression – PLS regression (BLANK e BROWN, 1993a).

O PCR é um método estatístico de análise multivariada que faz uso das componentes principais para a formação do modelo, sendo este em geral linear ou quadrático. Tanto o método PCR como o PLS são baseados no critério de minimização de uma função erro quadrático, similarmente às RNAs.

As RNAs podem representar um modelo de PCR quando as funções de transferências são lineares. Os pesos entre as camadas de entrada e intermediária são equivalentes aos *loadings* dos diferentes fatores e a ativação produzida pelos neurônios intermediários podem ser comparados com os *scores* do PCR.

A diferença entre o PCR e as RNAs está na maneira em que os parâmetros são otimizados. Em RNAs os parâmetros são ajustados sem a restrição de minimizar os resíduos quadráticos das amostras já em PCR a maximização da matriz de variância de X são também levados em consideração e portanto os parâmetros obtidos são diferentes.

Embora PCR seja um método linear, ele pode ser usado na modelagem de alguns tipos específicos de dados não lineares, tal como quadrático. E ainda se a forma de relação não linear entre as variáveis de entrada e aquela a ser predita é conhecida, um modelo pode ser linearizado pela transformação apropriada das variáveis originais, ou adicionando uma ordem maior e termos cruzados para a equação de regressão (DESPAGNE e MASSART, 1998). Deve-se entretanto, notar uma maior dificuldade na interpretação física de modelos

de ordem superior a dois, devido principalmente a existência de termos cruzados das componentes principais, que já são combinações lineares das variáveis originais.

A principal limitação do método PCR está em ser baseado na suposição de que uma relação simples (linear ou quadrática) existe entre a resposta do modelo e as componentes, e/ou na necessidade do conhecimento da forma desta relação se esta é não linear e de complexidade superior à quadrática.

2.4.5 Utilização da técnica de PCA no pré-processamento dos *inputs* das RNAs

Além de concentrar a maior parte da variância da matriz X em poucos eixos (as CPs), a técnica de PCA tem a propriedade de ortogonalizar estes eixos, que permite uma inversão mais precisa da matriz de variância-covariância na regressão de modelos lineares. Já para as RNAs a transformação ortogonal das variáveis eliminam problemas de modelagem ocasionados pela existência de correlação linear das variáveis de entrada da rede, auxiliando no mapeamento não linear da RNA e na otimização do seu tamanho (KANJILAL, 1995 e KOMPANY-ZAREH, 1999).

A razão entre o número de amostras e o número de parâmetros ajustáveis em uma RNA deve ser tão grande quanto possível. Assim, uma forma de se evitar o sobre ajuste do problema e conseqüentemente a redução do tamanho do conjunto de dados está na compressão dos mesmos (ZUPAN e GASTEIGER, 1993). A compressão permite ainda a eliminação de informações irrelevantes, como os ruídos e redundâncias presentes na matriz de dados.

O sucesso na compressão de dados pode acarretar num aumento da velocidade de treinamento, redução da memória de armazenamento, melhor habilidade de generalização do modelo, menor sensibilidade a ruídos nas medidas e representação de um modelo mais simples.

Nos trabalhos de CANCELLA e FANG (1996), DUTTA et al. (1998), HOLCOM et al. (1992) e KOMPANY-ZARED et al. (1999), entre outros, pode-se observar a relevância do uso da técnica de componentes principais no pré-processamento de dados de sistemas multivariados.

3 MATERIAIS E MÉTODOS

Este trabalho foi desenvolvido no Laboratório de Modelagem e Simulação de Processos Químicos - LMSPQ da Faculdade de Engenharia Química da UNICAMP. Para tanto foram utilizadas as ferramentas computacionais e metodologias apresentadas a seguir.

3.1 FERRAMENTAS COMPUTACIONAIS

Foram utilizados os seguintes pacotes computacionais:

- *Statistica*, para o pré-processamento dos dados e modelagem via PCR e
- *NeuroSolutions Educator*, para a aplicação das técnicas de Redes Neurais Artificiais na modelagem.

3.2 METODOLOGIA

A seguir é apresentado de forma clara e sucinta a metodologia desenvolvida neste trabalho.

3.2.1 Pesquisa bibliográfica

Nesta etapa foram levantadas as informações necessárias ao conhecimento das técnicas de modelagem e pré-processamento de dados citadas. Também foi realizado um estudo detalhado do processo de tratamento de efluentes da RIPASA S/A, incluindo sua dinâmica e a periodicidade de análise das principais variáveis de controle, e ainda os demais fatores que pudessem auxiliar a formulação do problema (tal como a produção de papel e celulose).

3.2.1.1 A RIPASA S/A

A RIPASA S/A – Indústria de Papel e Celulose localiza-se em Limeira, próxima ao rio Piracicaba, no estado de São Paulo. A RIPASA produz cerca de 750 tons/dia de celulose, sendo 95% branqueada e 650 tons/dia de papel. A madeira usada vem de árvores do gênero *Eucalyptus* proveniente das diversas plantações do grupo RIPASA. O papel é obtido pelo uso do processo *kraft* sendo adotada a seguinte sequência para o seu branqueamento: cloração, extração oxidativa com peróxido, hipocloração e peroxidação

Dos efluentes gerados pelo seu processo de produção aproximadamente 25% provêm da máquina de papel, 70% do branqueamento da celulose, e 5% da fábrica de recuperação.

Todos os efluentes de processo coletados são levados para o sistema de tratamento que consiste de um tratamento primário seguido de um tratamento biológico em lagoas aeradas.

Apesar da característica ácida dos efluentes do branqueamento, o efluente do lavador de gases do forno de cal permite que o efluente bruto tenha uma correção adequada do pH, evitando assim a necessidade de dosar corretivo.

Na estação de tratamento, o efluente passa por um decantador primário com capacidade de 6400 m³ para que os sólidos sedimentáveis presentes sejam separados (veja Fig. 3.1). Após este decantador, inicia-se o tratamento do efluente orgânico em duas lagoas aeradas, sendo que na primeira utiliza-se como complemento a injeção de oxigênio e adição de nutrientes.

A primeira lagoa possui 160.000 m³ e nela ocorre a maior redução de DBO. Em média a DBO inicial do efluente é de 550 mg/L, sendo reduzida para aproximadamente 100 mg/L. Esta lagoa possui 48 aeradores homogeneamente distribuídos e são injetados em média 120 m³/L de oxigênio puro. É feita ainda uma dosagem contínua de nutrientes (N, P) a fim de proporcionar condições adequadas de crescimento microbiológico e consequente redução da carga orgânica. Com o desenvolvimento dos microrganismos há um aumento da biomassa pela formação do lodo, o qual fica suspenso devido a aeração e deve ser retirado.

Assim utiliza-se um decantador secundário (6400 m³), o qual possui uma bomba que percorre toda sua extensão succionando o material decantado o qual juntamente com o material proveniente do primeiro decantador é destinado a lagoa de lama (10.000 m³),

seguindo para as valas de secagem de onde é retirado após algum tempo e usado como fertilizante em fazendas da região, conforme mostrado na Figura 3.1.

Posteriormente, o efluente é conduzido para a segunda lagoa (250.000 m³), onde há a redução de DBO de acordo com os padrões exigidos pela legislação, podendo então ser lançado ao rio Piracicaba.

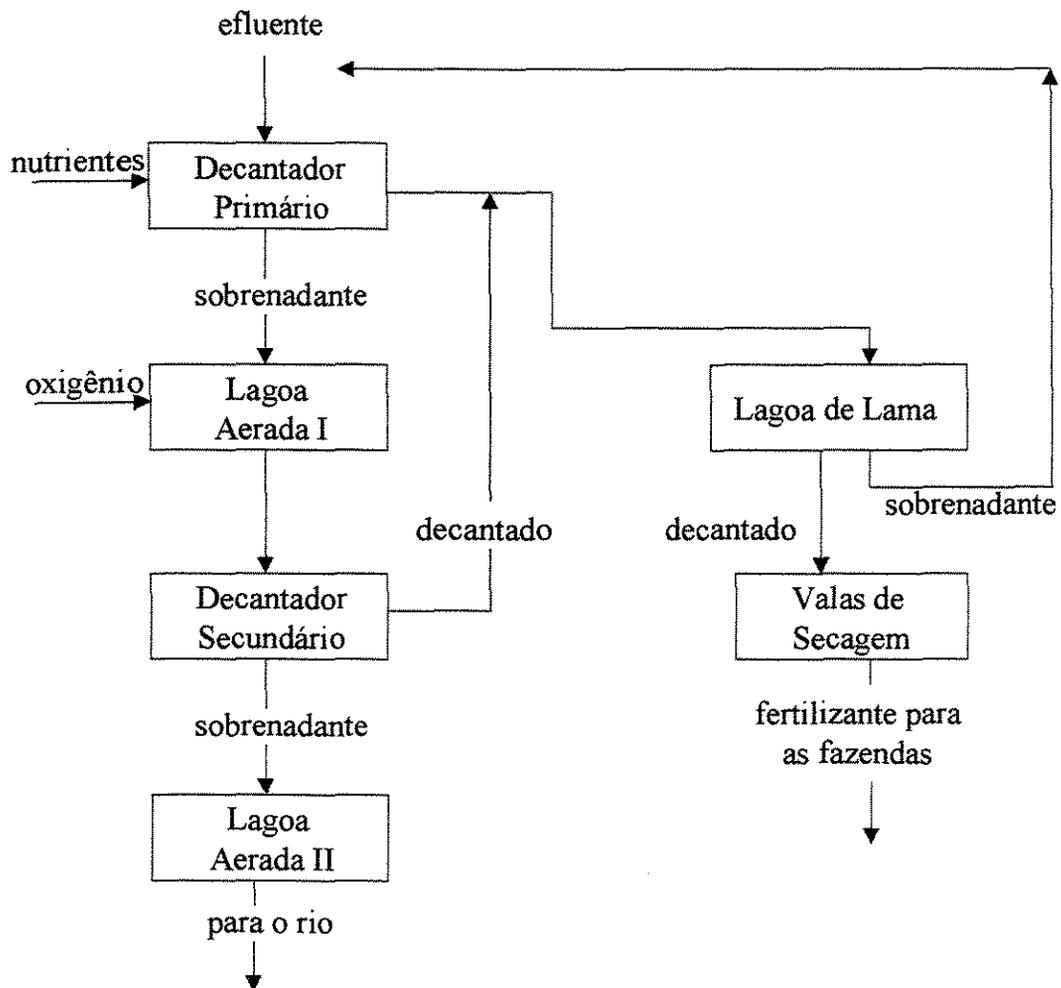


Figura 3.1: Sistema de tratamento de efluentes da RIPASA S.A.

Para o controle do seu sistema de lagoas aeradas, a RIPASA monitora os seguintes parâmetros: demanda química e bioquímica de oxigênio, vazão, pH e sólidos suspensos, apresentando apenas o pH medidas *on-line*.

3.2.2 Pré-processamento dos dados

Inicialmente foi necessária a organização dos dados fornecidos pela RIPASA, sendo estes referentes ao período de maio a novembro de 1997 e de janeiro a outubro de 1998.

Para a organização do banco de dados necessitou-se definir os parâmetros a serem utilizados no estudo, inclusive o tempo de residência do sistema de lagoas aeradas. Inicialmente foram considerados neste trabalho os parâmetros: DBO, DQO, vazão, pH, sólidos suspensos e produção de papel e celulose, sendo os três primeiros obtidos na entrada da lagoa aerada I e na saída da lagoa aerada II da RIPASA S/A.

3.2.2.1 Determinação do tempo de residência

O tempo de residência não foi um dado fornecido pela empresa, desta forma, sabendo-se que o tempo de residência médio de um efluente em uma lagoa aerada, ou sistemas de lagoas aeradas, regula entre 2 e 10 dias, os dados foram correlacionados em períodos de 3, 4, 5, 6 e 7 dias, para uma tentativa de sua determinação.

Para a determinação do tempo de residência realizou-se o mesmo procedimento do trabalho de GONTARSKI et al. (1998), ou seja, foram construídos gráficos do tipo “tempo *versus* valores de entrada e saída do sistema de lagoas” pois, uma vez que considera-se que os parâmetros de entrada determinam em parte como se comportarão os de saída, espera-se que o formato geral das curvas possam ser superpostas. Observa-se então a necessidade de um conjunto completo de dados de maneira a poder-se correlacionar as variáveis de entrada e saída assim como estas e a variável a ser predita (DBO_{out}).

3.2.2.2 Detecção de *outliers*

A detecção de *outlier* (JACKSON, 1991) apresenta dois passos: primeiramente, detecção de objetos atípicos, seguido pela identificação de *outliers*. Embora métodos numéricos, como a medida total da variabilidade - T^2 , permitam detectar as amostras que estão fora de controle, a identificação positiva do objeto atípico requer conhecimento do processo e procedimento de aquisição de dados.

Neste trabalho, apesar de ter-se identificado os objetos atípicos pelo estudo do processo e aquisição de dados, nenhum destes foi excluído do problema. Isto se deve ao fato do objetivo principal da modelagem do sistema de lagoas está em poder prever-se a DBO, ou a sua carga, principalmente quando esta apresenta-se fora dos seus valores pré-estabelecidos de controle. Além disso, havia a necessidade de trabalhar-se com todo o conjunto de dados devido à disponibilidade de um número reduzido destes.

3.2.2.3 Determinação dos parâmetros a serem utilizados para a predição da DBO_{out}

Devido a existência de lacunas, uma grande quantidade de dados teve que ser descartada, inclusive os parâmetros pH e os sólidos suspensos. A constância dos valores do pH frente às variações dos demais, inclusive à DBO_{out} , contribuiu também para a sua exclusão. Com esta manipulação de dados formou-se o primeiro grupo de estudo composto pelos parâmetros: DBO, DQO, vazão, e produção de papel e celulose.

Um segundo grupo foi formado substituindo-se a DBO e a DQO pelas suas cargas, isto é, pelos parâmetros de interação destes e da vazão. Desta forma nas modelagens trabalhou-se com a carga de DBO (CDBO), carga de DQO (CDQO) e vazão e produção de papel e celulose. A formação deste grupo foi estimulada tendo em vista o trabalho de GONTARSKI et al. (1999), que obteve resultados satisfatórios na modelagem de um sistema de tratamento quando utilizado um parâmetro de carga ao invés da DBO e DQO.

Através da técnica de PCA os parâmetros do segundo grupo foram ortogonalizados, sendo obtidas sete componentes principais. Assim quatro novos grupos foram formados pela combinação destas componentes principais, como mostrado na Tabela 3.2 Além da formação dos grupos destinados à modelagem, com o auxílio desta técnica, também estudou-se a relação entre os parâmetros mencionados.

Tabela 3.2: Grupos de variáveis utilizados na predição da DBO da saída da lagoa aerada II.

Grupos	Variáveis originais	Variáveis pré-processadas	Componentes principais
I	DQO _{in}		
	DQO _{out}		
	Vasão _{in} (Vas _{in})		
	Vasão _{out} (Vas _{out})	-	-
	Produção de Papel (Papel)		
	Produção de Celulose (Celulose)		
	DBO _{in} (DBO _{in})		
II		Carga de DQO _{in} (CDQO _{in})	
		Carga de DQO _{out} (CDQO _{out})	
		Vasão _{in} (Vas _{in})	
		Vasão _{out} (Vas _{out})	-
		Produção de Papel (Papel)	
	Produção de Celulose (Celulose)		
		Carga de DBO _{in} (CDBO _{in})	
III	-	-	Todos as componentes principais
IV	-	-	1°, 2°, 3°, 4°, 5° e 6° CPs
V	-	-	1°, 2°, 3°, 4°, 5° e 7° CPs
VI	-	-	1°, 2°, 3°, 4°, 5° CPs

in e out correspondem à entrada e saída do sistema de lagoas aeradas, respectivamente.

3.2.3 Desenvolvimento do modelo

Nesta etapa foram desenvolvidos modelos para predição da carga de DBO na saída da lagoa aerada II sendo estudadas as seguintes técnicas:

- PCA, para o pré-processamento de dados;
- RNA e PCR, para a obtenção do modelo.

3.2.3.1 Número de amostras e parâmetros

Mais que em qualquer método de regressão, o número de amostras é um fator limitante para o desenvolvimento de um modelo de RNA, tendo em vista a sua característica de “aprendizagem” através de exemplos.

Em geral o número de amostras é geralmente imposto ou limitado em problemas práticos, como neste estudo. Assim, como mencionado anteriormente, trabalhou-se com todos os dados disponíveis de maneira a manter a razão entre o número de amostras e o número de parâmetros ajustáveis tão alto quanto possível. Com isso tentava-se cobrir todo o domínio do problema e evitar sobre ajuste dos parâmetros e conseqüentemente perda de generalidade da rede, que pode ocorrer se poucos dados estão disponíveis.

3.2.3.2 Divisão de dados para treinamento, validação e teste

Para as RNAs o problema de modelagem é complexo pois com o ajuste dos dados de treino, deve-se verificar se o número de neurônios e camadas intermediárias é suficiente, buscando minimizar o erro da predição, o número de parâmetros a serem ajustados e o tempo de treinamento. Portanto, um conjunto de monitoramento adicional, validação ou teste, é necessário para parar o treinamento da rede antes de seu sobre ajuste. O conjunto de validação é apresentado à rede durante o treinamento e portanto, auxilia na decisão de quando parar o treinamento. Já o conjunto de teste é apresentado à rede já treinada. Como no caso da RIPASA a quantidade de dados disponíveis é pequena, trabalhou-se apenas com os conjuntos de treinamento e de validação (neste trabalho este último é denominado de teste por conveniência).

3.2.3.3 Determinação da topologia da rede

Diversas topologias da RNA têm sido propostas (HOPFIELD, 1982; FELDMAN e BALLARD, 1982; KOHONEN, 1984; RUMELHART e MCCLELLAND, 1986). Cada uma se diferencia pelo número e características dos neurônios, conexões, procedimentos de treinamento e se os valores de entrada e saída são discretos ou contínuos.

Neste trabalho utilizou-se uma rede *backpropagation feedforward*, com uma camada intermediária, fazendo-se variar o número de neurônios intermediários de 1 a 10, sendo testados os algoritmos de treinamento *Delta-Bar-Delta* (DBD) e *Quickpropagation* (QP).

A otimização da topologia das RNAs é provavelmente o passo mais tedioso do desenvolvimento do modelo, visto que o seu mapeamento não linear na busca do erro global mínimo geralmente é demorado, sendo encontrados apenas mínimos locais em alguns casos.

Número de camadas e neurônios

Foi trabalhado com uma única camada intermediária, tendo-se em vista que na modelagem de processos químicos o uso de uma única camada interna tem-se mostrado suficiente. Quando há a necessidade de modelos mais complexos o ajuste do número de neurônios da camada intermediária geralmente é suficiente.

Inicialmente foi trabalhado com sete neurônios de entrada da RNA, correspondentes apenas ao grupo II. Como o número de neurônios intermediários depende da complexidade das relações a serem aprendidas, o treinamento foi iniciado com uma unidade intermediária sendo este número aumentado gradativamente até dez.

Na tentativa de evitar-se um sobre dimensionamento da RNA, utilizou-se a técnica de PCA para minimização do número de neurônios de *input*. Através desta, os parâmetros de entrada da RNA foram substituídos pelas CPs e a seleção destas foi realizada com base nos autovalores e na porcentagem de variância acumulada explicada pelas componentes. Na Tabela 3.2 são apresentados os grupos formados com as CPs selecionadas.

Como para a camada de saída o número de neurônios corresponde ao número de variáveis a serem preditas, foi utilizado apenas um neurônio referente à $CDBO_{out}$.

Escolha da função de transferência

As funções de transferência geralmente utilizadas são a sigmoideal e a tangente hiperbólica que são facilmente diferenciáveis e permitem o ajuste dos mais variados modelos não lineares. Por apresentarem um comportamento aproximadamente linear no seu centro, estas tornam possível também a modelagem de sistemas lineares. A função sigmoideal foi então a escolhida como função de transferência tanto dos neurônios intermediários como no de saída.

3.2.3.4 Treinamento e Teste da rede

O problema do treinamento da rede corresponde basicamente à busca do mínimo global da função erro, Equação 5 (BISHOP, 1994). Este deve ser interrompido quando a rede apresentar uma boa capacidade de generalização e quando a taxa de erro for suficientemente pequena, ou seja, menor que um erro admissível. Tal momento não apresenta-se facilmente visível visto que a rede pode se especializar-se no conjunto de treinamento causando o problema de sobre ajuste (*over-training*).

Deste modo, a decisão de quando parar o treinamento da rede foi baseada no erro do conjunto de validação. Um erro de 0,001 foi tomado como critério de parada. Daí trabalhou-se com as diferentes topologias de redes, anteriormente definidas, sendo estas testadas diversas vezes de maneira a tentar-se encontrar o mínimo global, ou ao menos, um mínimo local que atendesse ao critério de parada estabelecido.

3.2.3.5 Interpretação e integração do modelo

Finalmente, após o treinamento e avaliação da rede, ela poderia ser integrada ao sistema operacional de aplicação, ou seja, ao sistema de lagoas aeradas da RIPASA. Esta etapa ainda encontra-se em desenvolvimento.

4 RESULTADOS E DISCUSSÕES

4.1 DETERMINAÇÃO DO TEMPO DE RESIDÊNCIA

A partir de tratamento dos dados apresentados no grupo II, foi possível montar arquivos para três, quatro, cinco, seis ou sete dias de residência e a construção dos gráficos de “tempo *versus* valores de entrada e saída” para determinação do tempo de residência (Figura 4.1).

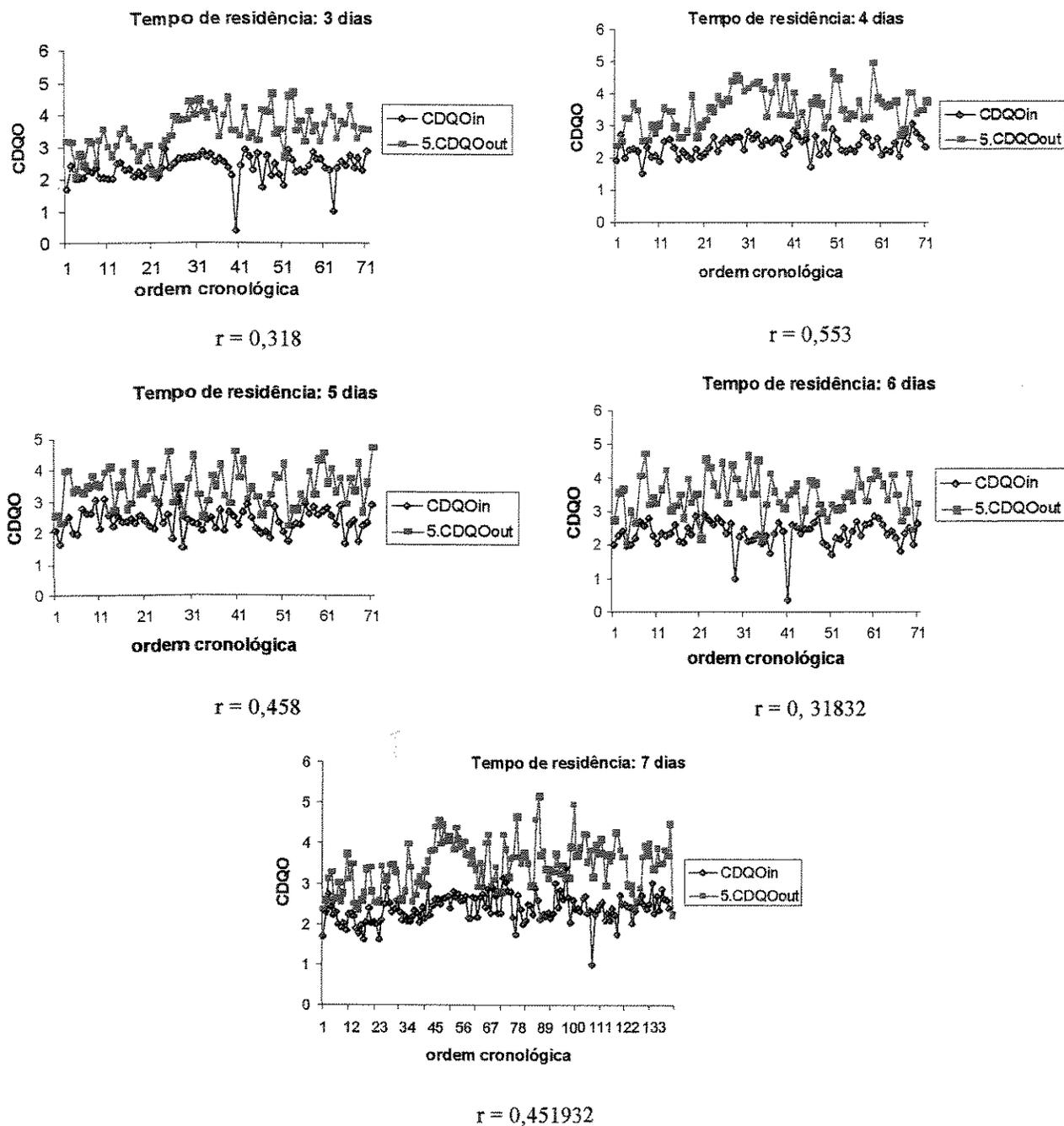


Figura 4.1: Gráficos comparativos entre DQO de entrada e DQO¹ de saída, e entre suas cargas, para os períodos de três a sete dias de residência.

¹ A escala dos valores de DQO de saída foram multiplicadas por 5 para que ambas as curvas fossem “plotadas” no mesmo gráfico.

O grande número de oscilações apresentados nestes gráficos e consequente incerteza quanto a melhor estimativa para o tempo de residência fez com que um banco de dados formados por todos os tempos inicialmente considerados (três a sete dias) fossem utilizados na modelagem por RNAs. Outro importante fator que contribuiu para que nenhum tempo de residência fosse excluído dos estudos consistiu na certeza de influência dos demais parâmetros sobre o valor da DQO de saída da lagoa II, afetando desta forma o perfil da relação entre este e a DQO de entrada.

4.2 RNA

Como mencionado na Seção 3.2.3.3, o treinamento da rede foi realizado utilizando-se o algoritmo padrão *backpropagation*, tendo como função de transferência a sigmoideal tanto na camada intermediária como na de saída.

Após a correlação dos parâmetros e formação dos conjuntos de dados do grupo II de acordo com os tempos de residência do sistema de lagoas considerados, 20 % destes foram destinados para o teste e o restante para o treinamento da rede, como apresentado na Tabela 4.1.

Tabela 4.1: Divisão de dados para composição dos arquivos de treino e teste².

Tempo de residência (dias)	Arquivo de treino	Arquivo de teste
3	57	14
4	57	14
5	75	19
6	57	14
7	112	28

Com testes para definição do algoritmo de treinamento (Delta-Bar-Delta ou Quickpropagation) e do número de neurônios da camada intermediária, montou-se os gráficos “erro de teste *versus* número de neurônios” para a predição da CDBO_{out}, apresentados na Figura 4.2.

² Os dados foram aleatorizados e 20 % destes compuseram o arquivo de teste e o restante o de treino. O número de dados de cada arquivo não é constante devido às lacunas existentes no banco de dados originais.

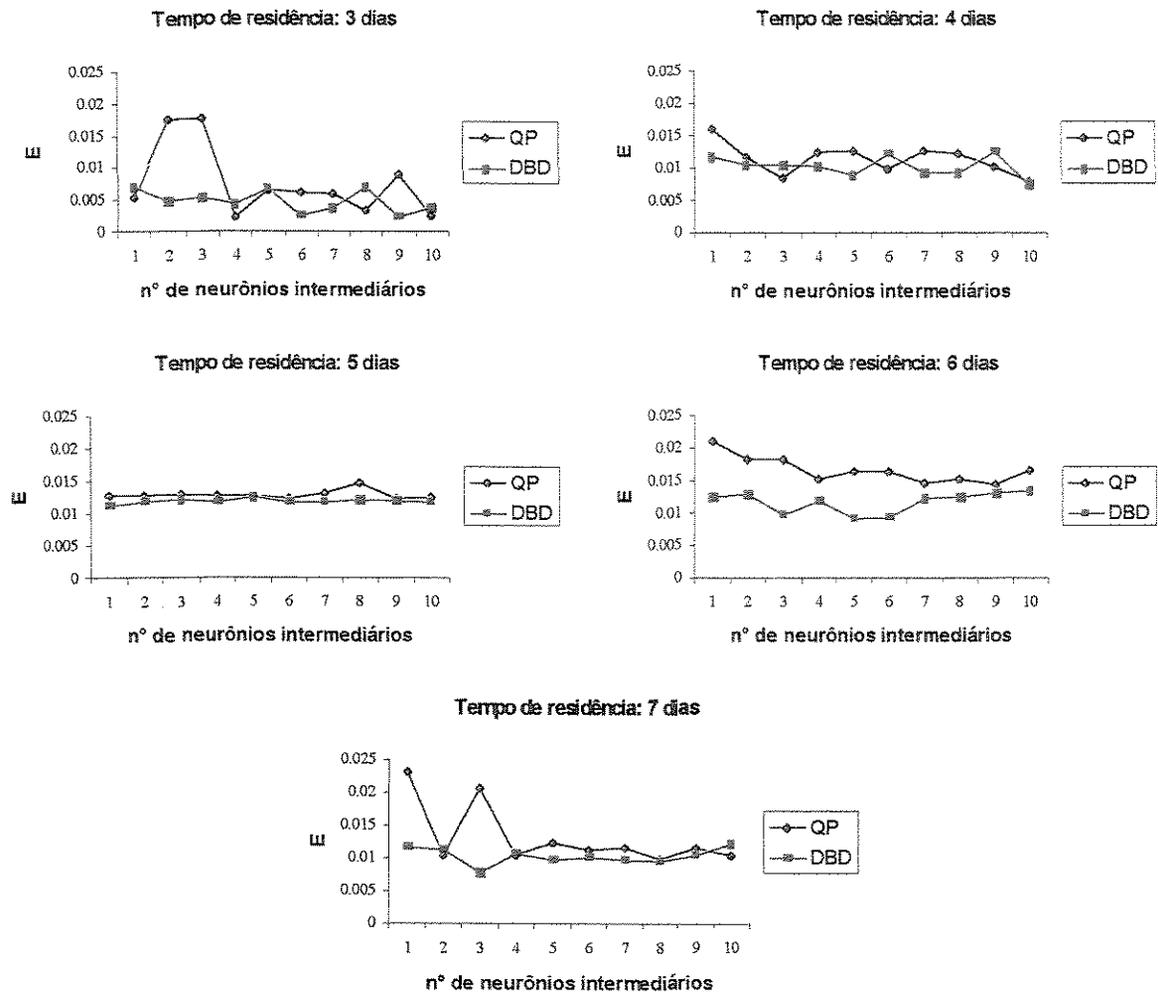


Figura 4.2: Gráficos da variação de E do conjunto de teste da $CDBO_{out}$ com o aumento do número de neurônios para os algoritmos Quickpropagation e Delta-Bar-Delta.

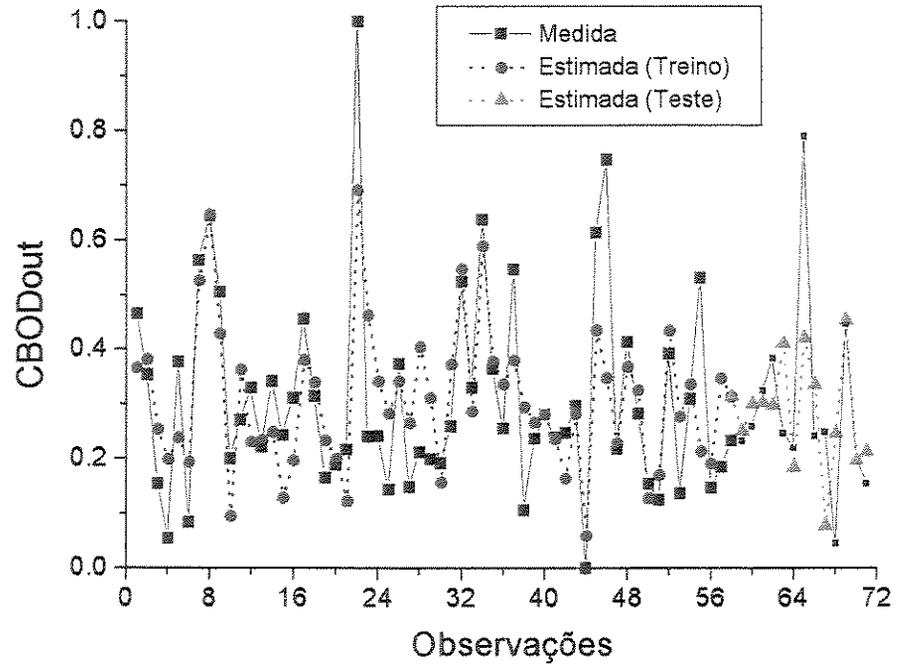
Com base no erro tanto do conjunto de teste como no de treino, foi realizada uma primeira triagem das redes testadas. Os três melhores resultados para cada tempo de residência considerado são apresentados na Tabela 4.2.

Tabela 4.2: Melhores resultados na predição da $CDBO_{out}$, considerando-se tempos de residência de 3 à 7 dias.

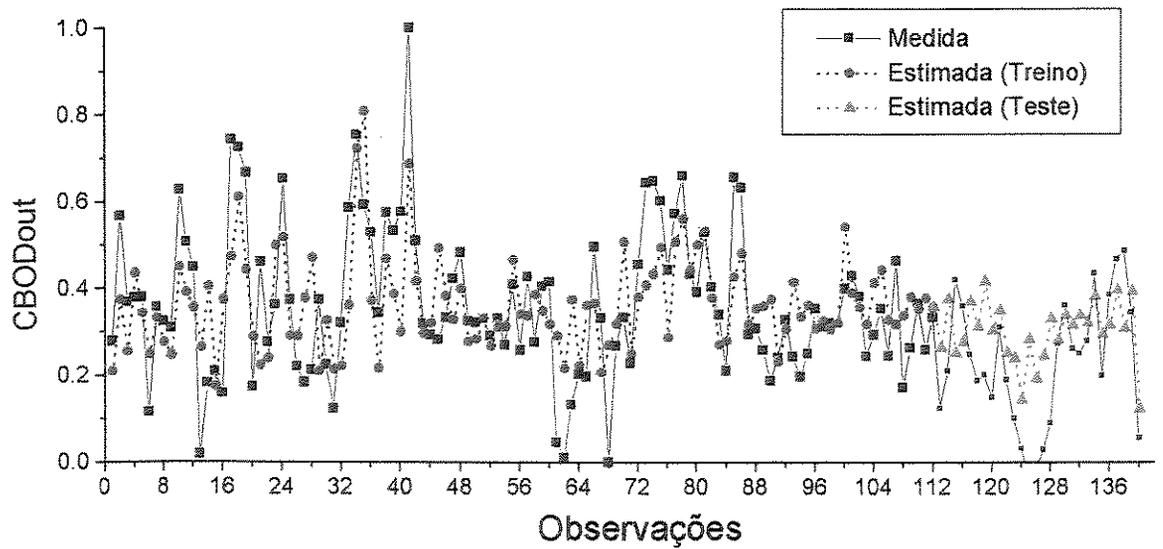
tempo de residência (dias)	Algoritmo	n° de neurônios intermediários	Erro quadrático médio - teste	Erro quadrático médio - treino
3	QP	4	0,0022	0,0258
	DBD	8	0,0023	0,0183
4	QP	10	0,0025	0,0246
	QP	3	0,0085	0,0287
	DBD	10	0,0075	0,0215
	QO	10	0,0080	0,0267
5	DBD	4	0,0118	0,0181
	DBD	6	0,0118	0,0175
	DBD	10	0,0117	0,0178
6	DBD	3	0,0097	0,0077
	DBD	5	0,0093	0,0075
	DBD	6	0,0094	0,0080
	DBD	5	0,0098	0,0089
7	DBD	7	0,0098	0,0091
	DBD	8	0,0095	0,0092

Da Figura 4.2 e Tabela 4.2, observa-se que a configuração interna definida pelo algoritmo Delta-Bar-Delta fornece um resultado geralmente superior ao do Quickpropagation.

Uma segunda triagem dos resultados foi realizada com base na necessidade de minimização dos parâmetros a serem ajustados pela rede, neste caso, escolha da rede foi baseada naquela que apresentou um menor número de neurônios intermediários. Desta forma, os melhores resultados, apresentados nas Figuras 4.3a e 4.3b, foram obtidos trabalhando-se com dados correlacionados para os tempos de seis e sete dias ambos com cinco neurônios.



(a)



(b)

Figura 4.3: Ajuste da rede para os dados da $CBOD_{out}$ do conjunto de treino e teste (tempo de residência de 6 e 7 dias)³.

Verificou-se que os resultados da modelagem utilizando as variáveis pré-processadas, ou seja CDQO e CDBO, foram consideravelmente melhores que os de GONTARSKI et al. (1998), veja Tabela 4.3 e APÊNDICE 1. Entretanto, o índice de correlação do conjunto de teste ainda continua baixo, significando que a rede é incapaz de prever com confiabilidade o comportamento de um novo conjunto de dados originários da mesma fonte que os de treinamento.

Tabela 4.3: Melhores resultados obtidos para a modelagem.

	Parâmetro predito	Tempo de residência	Índice de correlação (conjunto de teste)
GONTARSKI et.al. (1998)	DBO _{out}	5	0,42
		6	0,59
Estudo atual	CDBO _{out}	7	0,57

De maneira a conseguir-se melhorar os resultados e conseqüentemente integrar um modelo que apresente uma boa performance na predição da DBO_{out}, ou de sua carga, partiu-se para o estudo de uma nova técnica de pré-processamento de dados, Principal Component Analysis (PCA).

4.3 PCA e RNA

Dando-se continuidade ao estudo para a obtenção de um modelo de predição da CDBO_{out}, foi realizado um novo pré-processamento dos dados do grupo II utilizando-se a técnica de PCA. Neste foram admitidos apenas os dados correlacionados entre seis e sete dias por terem-se mostrado como melhores estimativas para o tempo de residência do sistema de lagoas aeradas da RIPASA.

Como descrito anteriormente, a técnica de PCA direciona as variáveis de entrada da rede de maneira a maximizar a variância dos dados. A análise da variância acumulada das componentes principais mostra que 94 e 91 % da variância nos dados de entrada pode ser

³ As observações não são apresentadas em ordem cronológica.

explicada pelas cinco primeiras componentes principais, para os tempos de residência de 6 e 7 dias, respectivamente (veja Tabela 4.4). Com base nessas variâncias foram formados os grupos III, IV V e VI apresentados na Tabela 3.1.

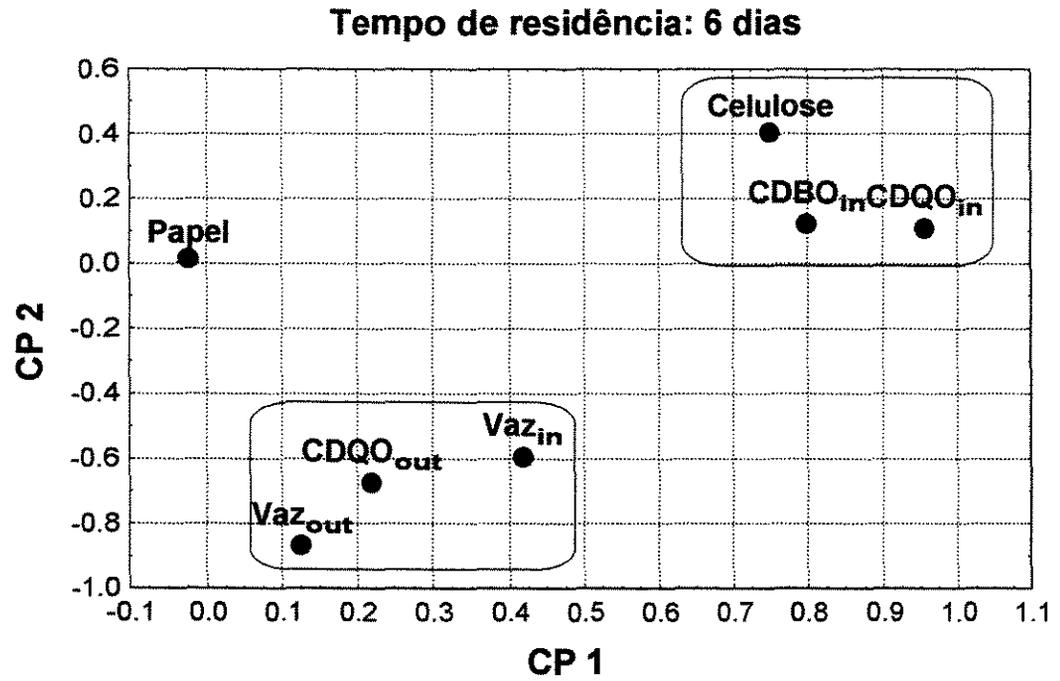
Tabela 4.4: Análise da variância acumulada das componentes principais.

Componentes principais	tempo de residência (dias)	
	6	7
	% de variância acumulada	
1°	32,9284	31,3990
2°	56,3688	52,9368
3°	72,4764	70,1006
4°	84,0855	83,1985
5°	93,2759	90,7862
6°	98,1272	97,0173
7°	100,0000	100,0000

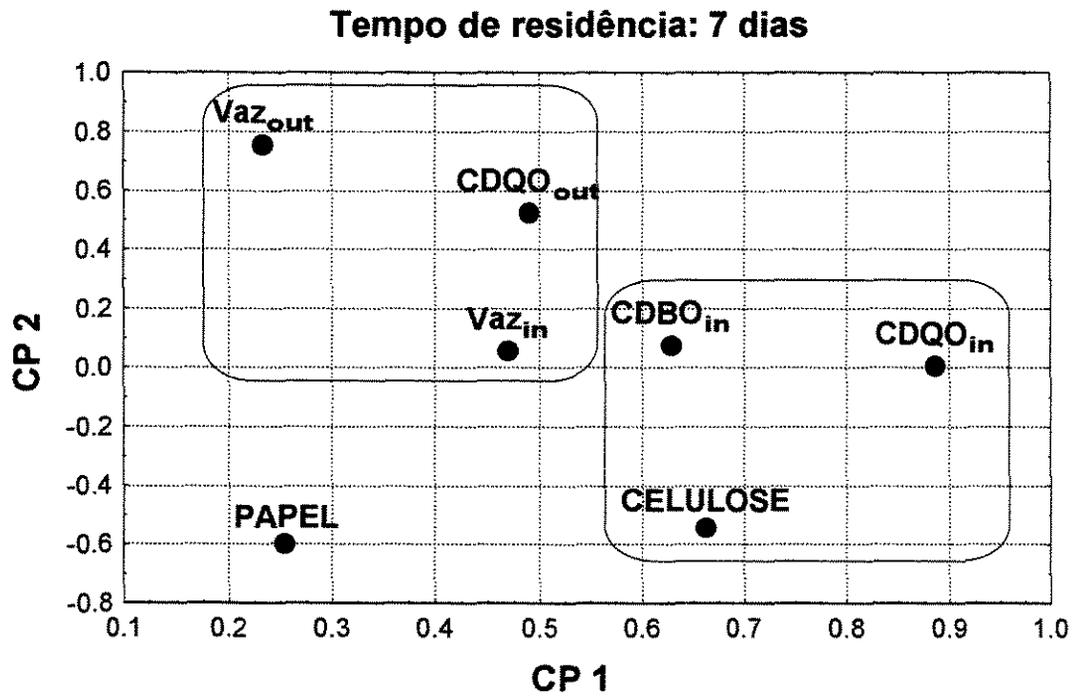
Plotando-se então os *loadings* das primeiras componentes principais, as quais explicam mais de 50 % da variância acumulada, é possível verificar se os parâmetros encontram-se correlacionados ou não.

Para o tempo de seis dias, dois *clusters* (grupos) distintos são observados, um associado com a produção de celulose, carga de DQO e DBO na entrada e o outro com a vazão na entrada e saída e carga de DQO na saída. A produção de papel aparece separada dos dois *clusters*. Este comportamento sugere uma pobre correlação entre o Papel e os demais parâmetros. Os mesmos agrupamentos são apresentados para o tempo de sete dias, entretanto os parâmetros encontram-se mais distribuídos indicando uma menor correlação entre eles (veja Figuras 4.4a e b).

Como no gráfico dos *loadings* apenas os parâmetros de *input* da rede são avaliados, não é possível ser verificada a existência de correlação entre estes e o parâmetro a ser predito, $CDBO_{out}$, portanto todos os parâmetros permaneceram nos estudos, inclusive a produção de papel.



(a)

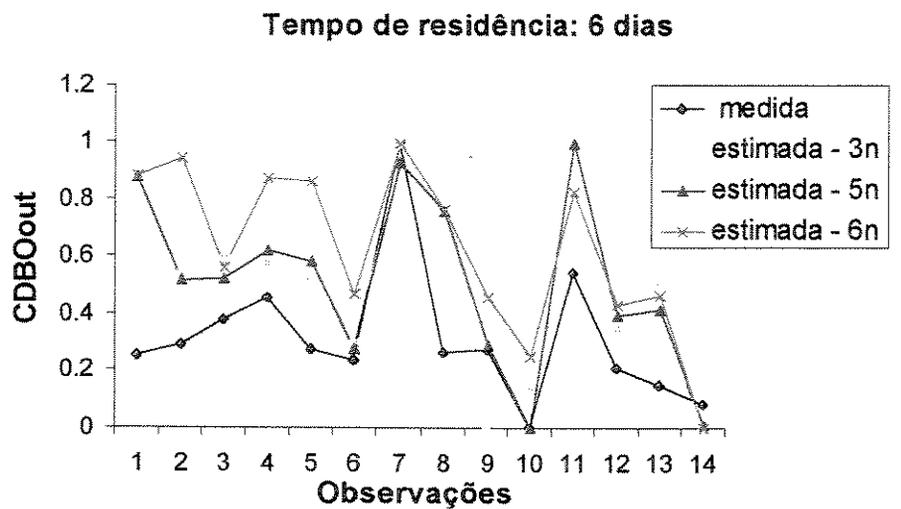


(b)

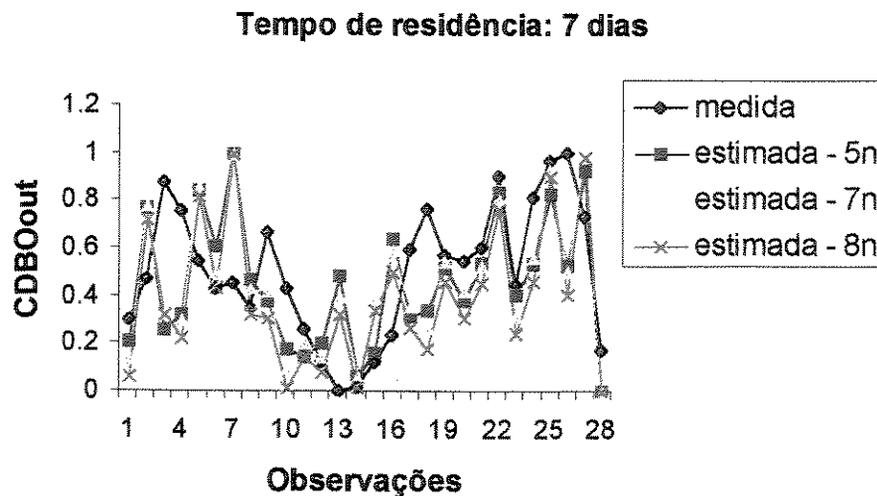
Figura 4.4: Gráficos dos loadings para os tempos de residência de 6 dias e 7 dias.

Após o pré-processamento dos dados por PCA, seguiu-se o mesmo procedimento da modelagem realizada para o grupo II. Inicialmente testou-se o grupo III composto pelos *scores* das sete componentes principais obtidas e pelos valores da carga de DBO de saída.

Na modelagem via PCA-RNA foram testadas apenas as topologias das redes que apresentaram os melhores resultados no estudo anterior, ou seja, algoritmo de treinamento Delta-Bar-Delta e número de neurônios intermediários de três, cinco e seis para o tempo de residência de seis dias e cinco, sete e oito para o tempo de sete dias. Os resultados são apresentados na Figura 4.5 e na Tabela 4.5.



(a)



(b)

Figura 4.5: CDBO_{out} medida e estimadas usando-se PCA-RNA dos conjuntos de teste correlacionados para os tempos de residência de 6 dias (a) e 7 dias (b).

Tabela 4.5: Resultados para a predição da $CDBO_{out}$ por PCA-RNA (algoritmo DBD).

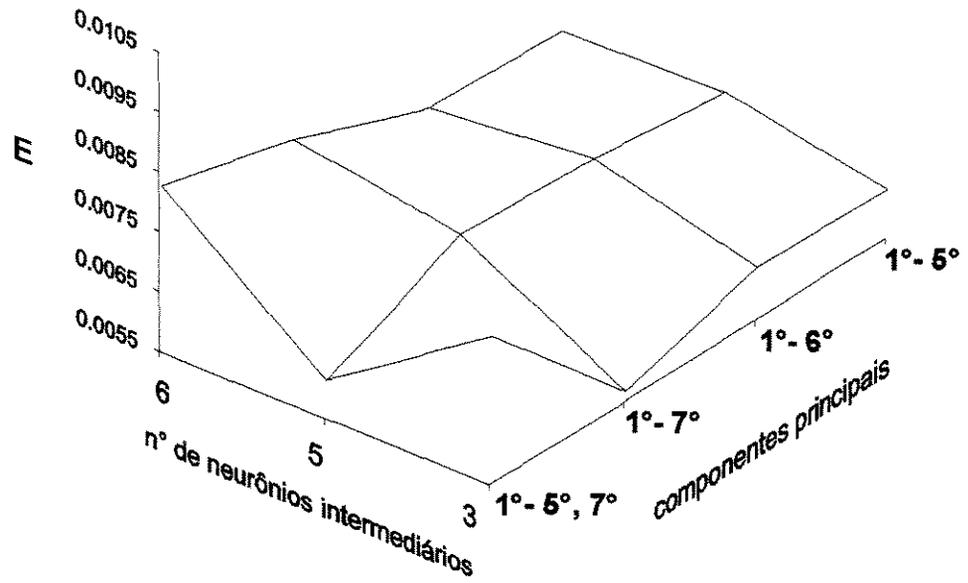
Arquivo	Medida da eficiência	Tempo de residência (dias)					
		6			7		
		n° de neurônios intermediários					
		3	5	6	5	7	8
treino	E	0,0057	0,0072	0,0077	0,0109	0,0100	0,0101
	r	0,81	0,76	0,74	0,63	0,62	0,62
teste	E	0,0082	0,0071	0,0084	0,0120	0,0112	0,0116
	r	0,68	0,72	0,67	0,58	0,57	0,59

onde: E e r correspondem ao erro quadrático médio e índice de correlação, respectivamente.

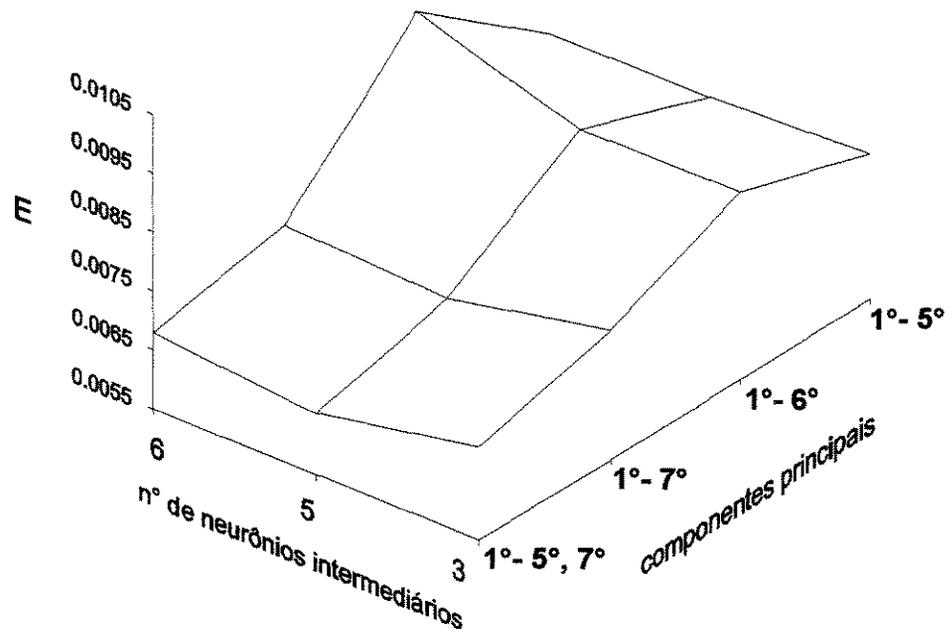
Verificou-se então que a melhor performance das redes testadas foi obtida quando utilizado cinco neurônios intermediários, para o tempo de seis dias. Partiu-se então para os testes dos grupos IV, V e VI para determinação do número apropriado de componentes principais. Nestes apenas trabalhou-se com os melhores resultados obtidos, ou seja, três, cinco e seis neurônios intermediários e o algoritmo de treinamento Delta-Bar-Delta, para um tempo de residência de 6 dias. Os resultados obtidos no treinamento e teste das redes encontram-se apresentados nas Tabela 4.6 e na Figura 4.6.

Tabela 4.6: Resultados para a predição da $CDBO_{out}$ utilizando as técnicas PCA-RNA (6dias).

Arquivo	componentes principais	Medidas da eficiência de predição	
		E	r
treino	1° ao 7°	0,0072	0,76
	1° ao 6°	0,0071	0,76
	1° ao 5°	0,0068	0,77
	1° a 5° e 7°	0,0061	0,80
teste	1° ao 7°	0,0071	0,72
	1° ao 6°	0,0086	0,67
	1° ao 5°	0,0078	0,70
	1° a 5° e 7°	0,0066	0,77



(a)



(b)

Figure 4.6: Erro quadrático médio vs nº de componentes principais vs nº número de neurônios intermediários para os conjuntos de treino (a) e teste (b).

Considerando-se todos os resultados obtidos por RNA e PCA-RNA, verificou-se que a melhor performance para predição da carga da DBO na saída da lagoa II é obtida quando os dados são pré-processados pela técnica de PCA antes de serem alimentados à RNA. Este resultado é obtido pela exclusão do sexto componente principal da camada de entrada para um rede composta de cinco unidades intermediárias, sendo utilizados os dados correlacionados para um tempo de 6 dias de residência. O ajuste dos dados para essa rede é apresentado na Figura 4.7 que segue.

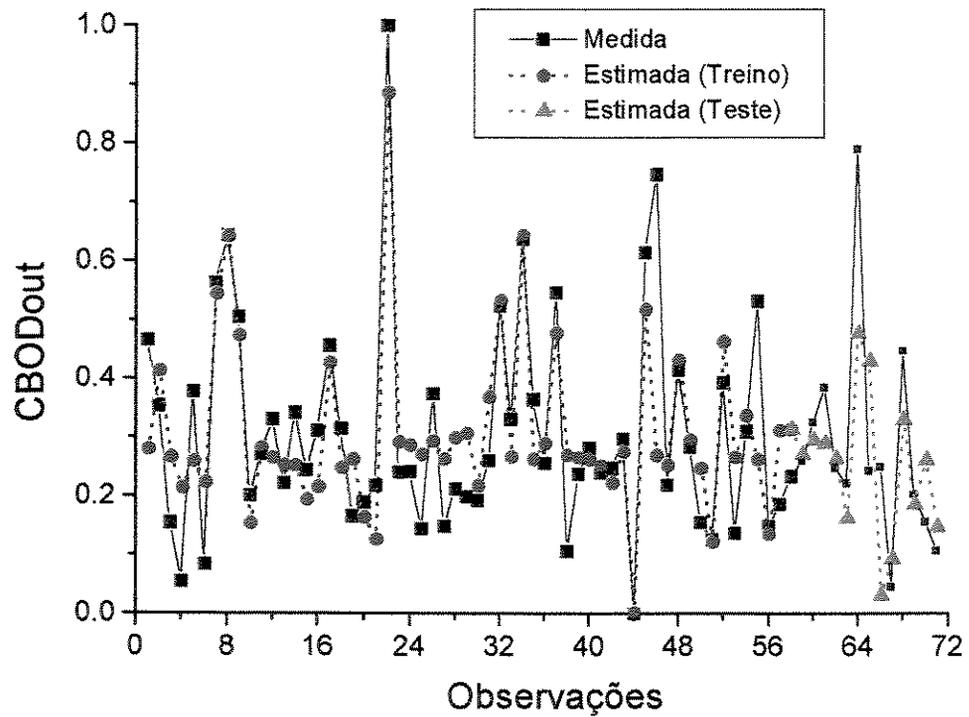


Figura 4.7: $CDBO_{out}$ medida vs estimada usando-se PCA-RNA (arquivos de treino e teste – 6 dias).

4.4 PCR

Na regressão das componentes principais (PCR) todos os *scores* foram usados, considerando a possibilidade que a variação (aproximadamente 6%) nos últimos PCs podem conter informações importantes para a predição da $CDBO_{out}$.

As Figuras 4.8 e 4.9 e a Tabela 4.7 mostram os melhores resultados para a predição pelas técnicas PCR^2 , RNA e RNA-PCA, onde PCR^2 indica uma regressão quadrática múltipla das CPs.

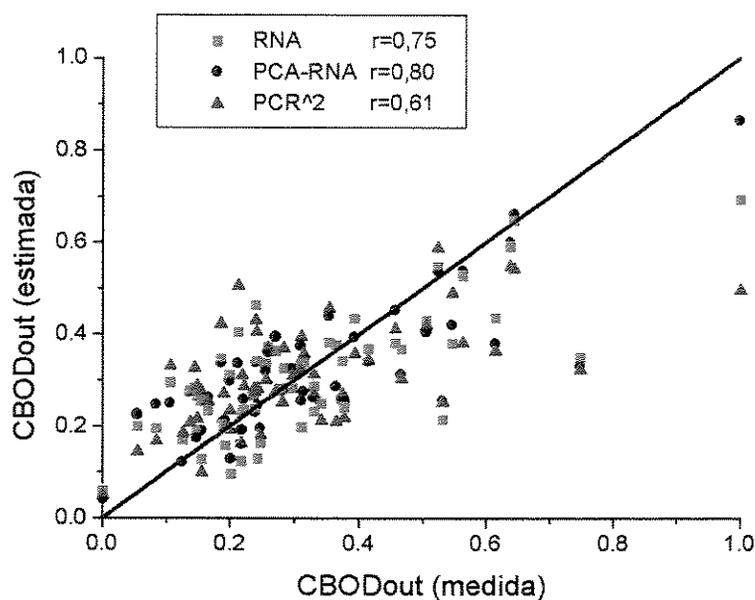


Figura 4.8: Correlação entre $CDBO_{out}$ estimada e medida pelos métodos RNA, PCA-RNA e PCR^2 (conjunto de treino)

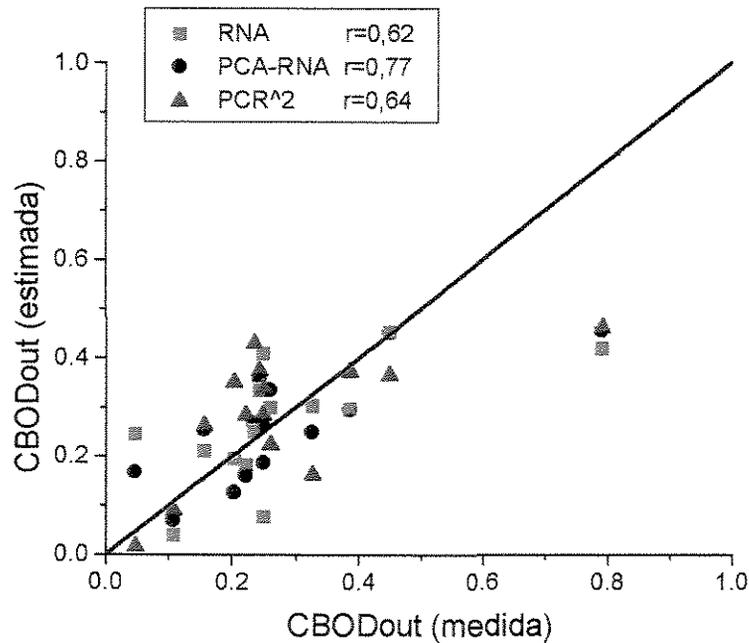


Figura 4.9: Correlação entre $CDBO_{out}$ estimada e medida pelos métodos RNA, PCA-RNA e PCR^2 (conjunto de teste)

Tabela 4.7: Resultados para a predição da $CDBO_{out}$.

Arquivo	estimativa da eficiência	RNA ^a	PCA-RNA ^b	PCR^2
treino	E	0,0077	0,0062	0,0105
treino	r	0,74	0,80	0,61
teste	E	0,0097	0,0065	0,0088
teste	r	0,60	0,77	0,64

^a rede 7-5-1 neurônios

^b rede 6-5-1 neurônios, sendo 1^o-5^o, 7^o CP como *inputs*

Através do método PCR^2 é obtido um resultado da predição superior ao da RNA para o conjunto de teste. Uma explicação para este fato pode consistir na existência de poucos dados resultando num sobre ajuste, *overfitting*, durante o treinamento da rede. O mesmo não ocorreu quando utilizadas as técnicas PCA e RNAs conjuntamente, as quais forneceram os melhores resultados de todas as predições.

Como mencionado, as principais vantagens para o tratamento dos dados pela técnica de PCA são a redução de dimensionalidade e a ortogonalidade. A transformação ortogonal pode levar à compactação de informação tornando possível a eliminação de dois problemas relacionados à predição via RNAs: excesso de parâmetros a serem ajustados e a correlação entre as variáveis de entrada. O problema de correlação em redes neurais tem sido estudado e reconhecido (DESPAGNE e MASSART, 1998; KANJILAL, 1995; KOMPANY-ZAREH *et al.*, 1999).

Neste trabalho a vantagem de redução da dimensionalidade oferecida pela técnica de PCA não parece ser tão importante embora uma melhora considerável dos resultados foi obtida pela exclusão do sexto componente principal das simulações. Este componente mostrou-se fisicamente insignificante e possivelmente associado a ruídos. Já o uso de variáveis ortogonais ajudou na otimização do tamanho da rede e na minimização da função erro.

5 CONCLUSÕES

As técnicas de PCA e RNAs foram aplicadas na modelagem de um sistema de lagoas aeradas. Várias arquiteturas de rede foram testadas em busca daquela que gerasse um menor erro para o conjunto de dados disponível.

Os altos erros encontrados para a modelagem via RNA mostram que as redes não foram satisfatoriamente treinadas e, portanto, não são capazes de predizer valores a partir de novos dados gerados na estação de tratamento.

Uma vez que as RNAs necessitam de um número considerável de dados, o motivo destes resultados pode estar na representatividade limitada dos dados disponíveis, que são descontínuos e correlacionados apenas por uma estimativa do tempo de residência

O resultado obtido pela regressão por PCR² apresentou-se superior ao obtido pela RNA, confirmando a dificuldade para formação de um modelo que a rede apresenta quando um número insuficiente de dados está disponível. Contudo, resultados satisfatórios da modelagem foram obtidos quando os dados foram pré-processados pela técnica de PCA antes de serem alimentados à rede.

Foi verificado que a técnica estatística de PCA é um importante pré-requisito para o pré-tratamento de dados. Esta possibilita a redução da dimensionalidade de um sistema a

ser modelado pela eliminação de variáveis que contêm ruído, além disso auxilia na redução do erro de modelagem via métodos estatísticos (como no caso do PCR²) e redes neurais.

Nesta aplicação, a redução de dimensionalidade obtida não foi significativa, mas em outras, que apresentam um número muito maior de variáveis de entrada, certamente a extração de fatores se faz necessária.

Apesar dos resultados apresentarem boa concordância com outros trabalhos, tais como os de BLAESI e JENSEN (1992) e HÄCK e KÖHNE (1996), a integração do modelo PCA-RNA para monitoração da demanda bioquímica de oxigênio do sistema de lagoas aeradas está sendo avaliado pela RIPASA S/A. O próximo passo será o teste deste modelo pela apresentação de um novo conjunto de dados RIPASA S/A.

6 SUGESTÕES

Imediatas

- Verificação da capacidade de generalização da rede através da apresentação de um novo conjunto de dados à mesma.

A curto prazo

- Adequação do sistema de fornecimento de dados para possibilitar a formação de um banco de dados consistente e sem lacunas.

A longo prazo

- Obtenção de um novo banco de dados;
- Nova modelagem via RNAs;
- Aplicação de técnicas de préprocessamento de dados, como PCA e PLS - Partial Least Square.

7 REFERÊNCIAS BIBLIOGRÁFICAS

- ADAMS, M.J., "Chemometrics in analytical spectroscopy", Cambridge, UK (1995).
- BISHOP, C.M., "Neural networks and their applications", *Rev. Sci. Instrum.*, v. 65, n. 6, 1803-1831 (1994).
- BLAESI, J. and B. JENSEN, "Can neural networks compete with process calculations?", *Intech Applying Technology*, 34-37 (1992).
- BLANK, T.B. and S.D. BROWN, "Data preprocessing using neural networks", *Anal. Chim.*, v. 277, 273-287 (1993a).
- BLANK, T.B. and S.D. BROWN, "Nonlinear multivariate mapping of chemical data using feed-forward neural networks", *Anal. Chem.*, v. 65, 3081-3089 (1993b).
- BOUGGAARD, C. and H.H. THODBERG, "Optimal minimal neural interpretation of spectra", *Anal. Chem.*, v. 64, 545-551 (1992).
- CANCILLA, D.A and X. FAND, "Evaluation and quality control of environment analytical data from the Niagara River using multiple chemometric method", *Journal Great Lakes Res.*, v. 22, 241-253 (1996).
- CONNELL, D. W. e G. J. MILLER, "Chemistry and ecotoxicology of pollution", New York: J. Wiley (1984).
- COTE, M., B.P.A. GRANDJEAN, P. LESSARD and J. THIBAU, "Dinamic modeling of activated sludge process: Improving prediction using neural networks", *Water Res.*, v. 29, n. 4, 995-1004 (1995).
- DESPAGNE F. and D. L. MASSART, "Neural networks in multivariate calibration", *Analyst*, 123, 157-178 (1998).
- DUTTA, D., A. K. MOHANTY, R.K. CHOUDHURY and P. CHAND, "Pattern recognition of particles tracks using principal component analysis and artificial neural network", *Nuclear Instruments & Methods in Physics Research*, 404, 445-454 (1998).
- FELDMAN, J.A and D.H. BALLARD, "Connectionist models and their properties", *Cognitive Sci.*, v. 6, 205-254 (1982).
- GALVÃO, J. B., V. M. GRIECO, R. P. ARAÚJO, M. R. ORTOLANO, E. BERTOLETTI e M. L. L. C. RAMOS, "Treatability Studies and Toxicity Reduction in Pulp and Paper Mill Effluents", *Water Science and Technology*, 20, 149-160, 1998.

GONTARSKI C.A., C. CASTRO, C. SOUZA, L. FERRARI, J. VITI, M. MORI e A.M. PEREIRA, "Avaliação do processo de tratamento de resíduos industriais de papel e celulose com redes neuronais", *XVIII Interamerican Congress of Chemical Engineering*, Puerto Rico (1998).

GONTARSKI C.A., W. BONIFÁCIO e M. MORI, "Predição de comportamento de um sistema de tratamento de efluentes industriais utilizando redes neuronais", *Anais do II Congresso de Engenharia de Processos do Mercosul – ENPROMER 99*, Florianópolis – SC (1999).

GONZÁLEZ, A.G. and D. GONZÁLEZ-ARJONA, "Statistical assessment of a new criterion for selecting the number of factors in factor analysis", *Anal. Chim.*, v. 314, 251-252 (1995).

HÄCK, M. and M. KÖHNE., "Estimation of wastewater process parameters using neural networks", *Water Sci. Technol.*, v. 33, n. 1, 101-107 (1996).

HAMMER, M.J., *Sistemas de abastecimento de água e esgotos*, Rio de Janeiro: Livros técnicos e Científicos Editora S.A. (1979).

HOLCOM T.R. and M. MORARI, "PLS/ Neural Networks", *Computers Chem. Eng.*, v. 16, n. 4, 393-411 (1992).

HOLPFIELD, J.J., "Neural networks and physical systems with emergent collective computational abilities", *in Proc. Nat. Academy Sci.*, v. 79, 2554-2558 (1982).

HOSKINS, J.C. and D.M. HIMMENBLAU, "Artificial neural network model of knowledge representation in chemical engineering", *Computers Chem. Eng.*, v. 12, 881-890 (1988).

HOSKINS, J.C. and D.M. HIMMENBLAU, "Process control via artificial neural networks and reinforcement learning", *Computers and Chem. Eng.*, v. 16, n. 4, 241-251 (1992).

JACKSON, J.E., "A user guide to principal components", Wiley Press, New York (1991).

JACOBS, R.A., "Increased rates of convergence through learning rate adaptation", *Neural Networks*, v. 1, 295-307 (1988).

KANJILAL P.P., "On the application of orthogonal transformation for design and analysis of feedforward networks", *IEEE Trans. Neural Networks*, v. 6, 1061-1070 (1995).

KARIM, M.N. and S.L. RIVERA, "Comparison of feed-forward and recurrent neural networks for bioprocess state estimation", *Suplemento da Computers and Chem. Eng.*, v. 16/S, 1369-377 (1992).

- KOHONEN, T., "Self-organization and associative memory", Springer-Verlag, Berlin (1984).
- KOMPANY-ZARED M., A. MASSOUMI e S.H. PEZESHK-ZADEH, "Simultaneous spectrophotometric determination of Fe and Ni with xylene orange using principal component analysis and artificial neural networks in some industrial samples", *Talanta*, 48, 283-292 (1999).
- MAJCEN, N., K. Rajer-Kanduc, M. Novic e J. Zupan, "Modeling of property prediction from multicomponent analytical data using different neural networks", *Anal. Chem.*, v. 67, 2154-2161 (1995).
- MARDIA, K. V., J. T. KENT e J. M. BIBBY, "Multivariate Analysis", Academic Press, 1979.
- METCALF & EDDY, "Wastewater engineering: treatment, disposal and reuse, 3.ed. New York: McGraw-Hill, 1991.
- MINSKY, M. e S. PAPERTS, "Perceptrons: an introduction to computational geometry, MIT Press, Cambridge, Mass., (1969).
- MONTEIRO, C.E, "Poluição das águas - autodepuração dos corpos d'água. São Paulo: CETESB / ABES / BNH", 36, (1975). (Apostila).
- MORRIS, A.J., G.A. MONTAGUE and M.J. WILLS, "Artificial neural networks: studies in process modeling and control", *Transactions of IChemE*, 72, A, 3-19 (1994).
- OLIVEIRA, K.P.S., M. MORI e R.E. BRUNS, "Industrial effluent treatment system modeling using PCA and neural networks", submetido à publicação em uma revista internacional indexada (1999).
- PSICHOGIOS, D.C. and L.H. UNGAR, "A hybrid neural network-first principles approach to process modeling", *AIChE Journal*, v. 38, 1499-1511 (1992).
- PIELOU, E.C., "The interpretation of ecological data", Wiley Press, New York, USA (1984).
- PONTON, J.W. and J. KLEMES, "Alternatives to neural networks for inferential measurement", *Computers Chem. Eng.*, v. 17, n. 10, 991-1000 (1993).

PU, H. and Y. HUNG, "Use of artificial neural networks: Predicting trickling filter performance in a municipal wastewater treatment plant", *Envir. Manag. Health*, v. 6, n. 2, p. 16-27 (1995).

ROSENBLATT, R., "Principles of Neurodynamics", Spartan Books, New York (1962).

RUMELHART, D.E. and J.L. McCLELLAND, editors, "Parallel distributed processing: explorations in the microstructure of cognition, v. 1, Foundations Mit Press (1986).

SAVKOVIC-STEVANOVIC, J., "Neural networks for process analysis and optimization: modeling and applications", *Computers Chem. Eng.*, v. 18, n. 11, 1149-1155 (1994).

YANG, G.B.O., "Managing secondary treatment systems", *TAPPI Journal*, v. 52-54 (1996).

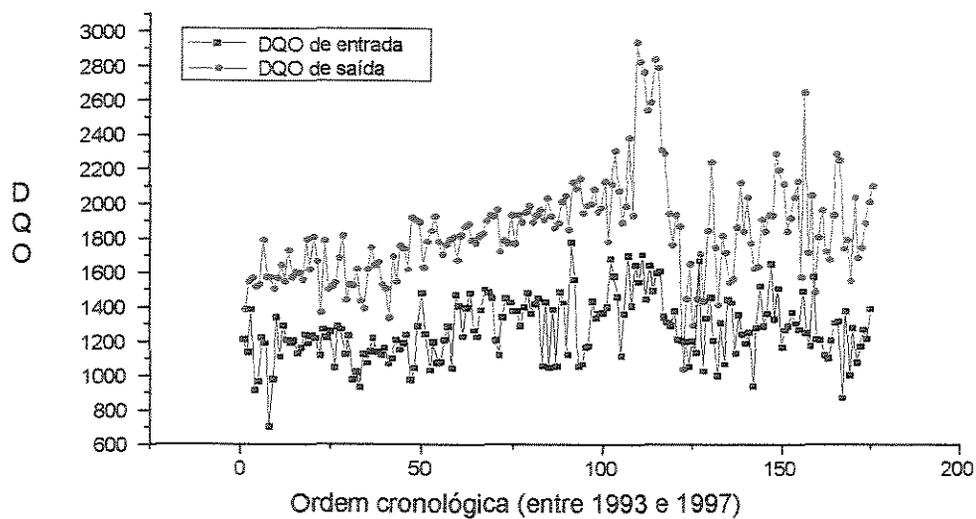
WALCZAK, B. and D.L. MASSART, "The Radial Basis Functions – Partial Least Squares approach as a flexible non-linear regression technique", *Anal. Chim.*, v. 331, 177-185 (1996).

ZUPAN, Z. and J. GASTEIGER, "Neural Networks for Chemists", New York, USA (1993).

APÊNDICE 1

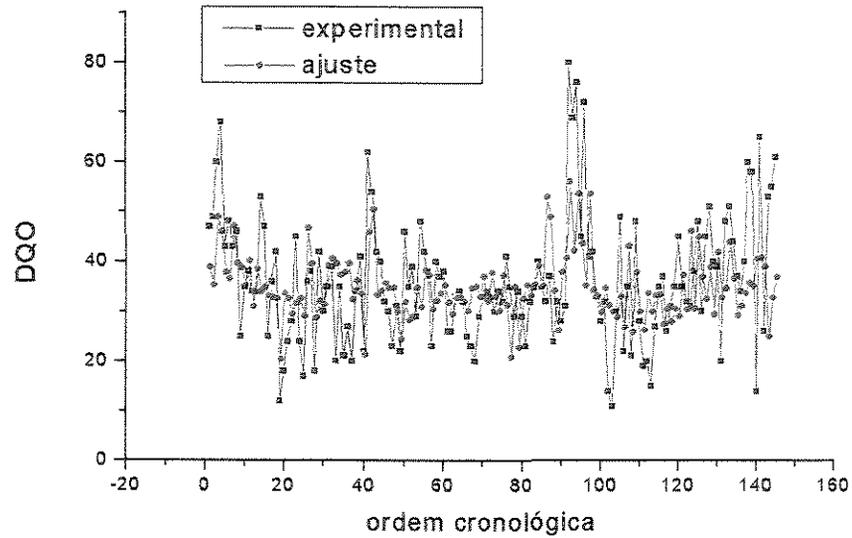
Resultados obtidos por (GONTARSKI et. al., 1998)

APÊNDICE 1A: Gráfico comparativo entre DQO de entrada e DQO¹ de saída para um período de cinco dias de residência.



¹ A escala dos valores de DQO de saída multiplicada por 4 para que ambas as curvas fossem "plotadas" no mesmo gráfico.

APÊNDICE 1B: Ajuste da rede para dados da DBO do conjunto de treino (tempo de residência de 5 dias).



APÊNDICE 1C: Ajuste da rede para dados da DBO do conjunto de teste (tempo de residência de 5 dias).

