



Trabalhos Técnicos

**19º CONGRESSO BRASILEIRO DE
ENGENHARIA SANITÁRIA E AMBIENTAL**

TEMA CENTRAL
SANEAMENTO AMBIENTAL:
AMBIENTE DE NEGÓCIOS OU UM NEGÓCIO AMBIENTAL?

**2ª FEIRA INTERNACIONAL DE TECNOLOGIAS
DE SANEAMENTO AMBIENTAL - FITABES'97**

14 A 19 DE SETEMBRO / 97

LOCAL:

Rafain
PALACE HOTEL

FOZ DO IGUAÇU - PR

PROMOÇÃO / REALIZAÇÃO:



ABES - Associação Brasileira de
Engenharia Sanitária e Ambiental
Capítulo Nacional da AIDIS

MODELAGEM DE UMA ESTAÇÃO DE TRATAMENTO DE ESGOTOS UTILIZANDO REDES NEURAIS: CONCEITUAÇÃO E AVALIAÇÃO PRELIMINAR

Werner Siegfried Hanisch⁽¹⁾

Engenheiro Químico pela Universidade Federal de São Carlos, Mestre em Hidráulica e Saneamento pela Escola de Engenharia de São Carlos da Universidade de São Paulo. Doutorando pelo Departamento de Hidráulica e Saneamento da Escola de Engenharia de São Carlos.

Claudia Regina Milaré

Graduada em Ciências da Computação pela Universidade Estadual de Londrina - UEL, Mestranda do Instituto de Ciências Matemáticas e Computacionais de São Carlos - USP na área de Ciências de Computação e Matemática Computacional.

André Carlos Ponce de Leon Ferreira de Carvalho

PhD em Engenharia Eletrônica em 1994 pela University of Kent at Canterbury. Professor MS-3 do Departamento de Ciências de Computação e Estatística da Universidade de São Paulo.

Eduardo Cleto Pires

Engenheiro mecânico pela Escola de Engenharia de São Carlos, EESC-USP; Mestre em Engenharia mecânica pela PUC-RJ e Doutor e Livre Docente em Hidráulica e Saneamento pela Escola de Engenharia de São Carlos, EESC-USP. Professor Associado do Departamento de Hidráulica e Saneamento da Escola de Engenharia de São Carlos, EESC-USP.



Endereço⁽¹⁾: Av. Dr. Carlos Botelho, 1465 - Centro - São Carlos - SP - CEP: 13560-250 - Brasil - Tel: (016) 274-9264 - Fax: (016) 274-9212 - e-mail: werner@icmssc.sc.usp.br.

RESUMO

Uma Estação de Tratamento de Esgotos (ETE) deve funcionar de maneira correta para não acarretar problemas tanto do ponto de vista social, quanto do ponto de vista biológico. Todos os resíduos provenientes das atividades humanas devem ser tratados antes de serem dispostos no ambiente. Algumas aplicações de problemas relacionados ao meio ambiente envolvem processos dinâmicos que são difíceis de serem descritos por métodos matemáticos convencionais. Além disso o controle destes processos apresenta algumas dificuldades, como por exemplo, a presença de sistemas não-lineares com vários graus de liberdade e incerteza. A modelagem de uma ETE é um problema com tais características. As aplicações que utilizam redes neurais aumentaram significativamente nos últimos anos em diversas áreas. Esta tecnologia se encaixa muito bem em problemas de controle, possibilitando um desempenho superior ao dos modelos convencionais. Este trabalho tem como objetivo principal iniciar os estudos de modelagem de uma ETE utilizando redes neurais, visando a previsão dos parâmetros de operação, apoiados em informações passadas. Os resultados obtidos indicam que redes neurais podem ser adequadas para modelar estações de tratamento de esgotos.

PALAVRAS-CHAVE: Redes Neurais, Modelos Matemáticos, Processos Aeróbios.

INTRODUÇÃO

O mau funcionamento de uma Estação de Tratamento de Esgotos (ETE) traz sérios problemas ambientais e sociais, visto que os efluentes destas estações podem causar ou alastrar diversas doenças aos seres humanos, assim como desestabilizar o meio natural no qual estes despejos são lançados (Belanche, 1992).

Ainda que conscientes da grande importância do tratamento de esgotos, no Brasil ainda muito pouco se tem investido para melhorar as redes coletoras de esgotos e posterior tratamento. Apenas 30% da população urbana brasileira têm seus esgotos coletados, e destes, apenas 10% têm tratamento adequado (Costa Júnior, 1996; Tinôco Filho, 1996). Entretanto, esforços têm sido concentrados para baratear e melhorar as tecnologias disponíveis no mercado para o tratamento de esgotos, que podem levar a uma ação mais efetiva por parte dos governantes na utilização destes sistemas para a melhoria da qualidade de vida da população brasileira.

Com o intuito de melhorar a operação destas estações de tratamento, a emissão de diagnósticos a partir das observações dos parâmetros na entrada, saída e etapas intermediárias do processo, possibilita um maior controle do processo de tratamento de esgoto. Para a resolução desta tarefa é necessário, inicialmente, determinar um modelo matemático que reproduza, com uma precisão especificada, o processo de tratamento como um todo, englobadas todas as operações unitárias e processos biológicos.

Devido à dificuldade de se modelar tais processos, caracterizados por serem altamente complexos e não-lineares, este trabalho propõe a aplicação de redes neurais artificiais para simular o comportamento de uma ETE convencional, visto que a forma como os problemas são modelados pela abordagem de redes neurais e o paralelismo natural inerente à sua arquitetura, criam a possibilidade de um desempenho superior à dos modelos convencionais.

O objetivo desta primeira etapa do trabalho é a elaboração de uma função que prevê os parâmetros de operação da ETE em um instante de tempo futuro, baseados em informações presentes e passadas. Os dados utilizados são de uma ETE em operação em Manresa, Espanha. Esta planta consiste de 6 bioreatores aeróbios de 1800m³ divididos em duas linhas de três reatores, dois decantadores primários de 37m de diâmetro e volume de 3332m³ e dois decantadores secundários de 42m de diâmetro e volume de 4850m³. A temperatura da água residuária oscila entre 12°C no inverno e 20°C no verão. A taxa de recirculação varia em torno de 100% da vazão de entrada e a idade do lodo está situada entre 4 e 7 dias (Poch, 1993).

No próximo item alguns conceitos básicos sobre redes neurais serão apresentados e literaturas adequadas ao seu pleno entendimento serão sugeridas. Logo após, a metodologia, resultados, discussões e conclusões serão descritos.

REDES NEURAIS

Redes neurais são sistemas paralelos distribuídos, compostos por unidades de processamento simples (neurônios). Estas unidades computam certas funções matemáticas e estão dispostas em uma ou mais camadas, interligadas por um grande número de conexões geralmente unidirecionais. Estas conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo, e servem para ponderar a entrada recebida em cada neurônio da rede. O funcionamento destas redes é inspirado no cérebro humano e têm como atribuições mais comuns a aproximação de funções lineares ou não-lineares, classificação e reconhecimento de padrões. Suas semelhanças com o cérebro humano baseiam-se em dois aspectos fundamentais, segundo Haykin (1994).

1. Conhecimento é adquirido pela rede através de um processo de aprendizado e;
2. A intensidade, ou força da conexão entre os neurônios, conhecida como pesos sinápticos são usadas para armazenar o conhecimento.

Existem diversas arquiteturas de redes, sendo que a rede do tipo MLP (perceptron de múltiplas camadas) tem sido ostensivamente estudada no campo da Engenharia para simulação de processos químicos e biológicos, pois podem aproximar qualquer tipo de função não-linear (Cruz e Giordano, 1995; Morris et al., 1994; Bhat et al., 1990). Assim, optou-se trabalhar com este tipo de rede para modelar a ETE. Esta rede é mostrada na Figura 1.

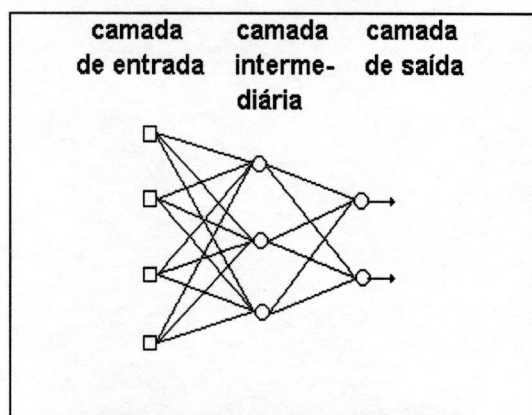


Figura 1: Redes MLP completamente conectadas

Fonte: Haykin (1994)

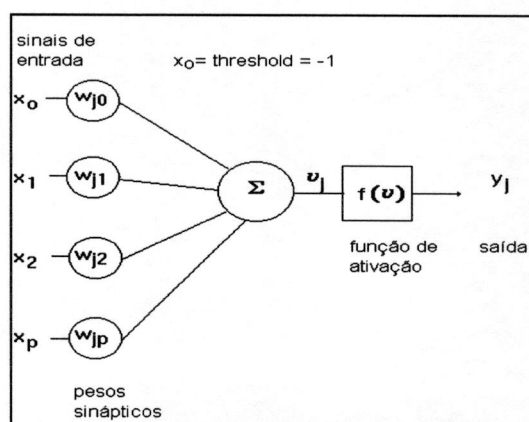


Figura 2: Modelo não-linear de um neurônio

Fonte: Haykin (1994)

As redes MLP são caracterizadas por possuírem uma ou mais camadas internas (escondidas), cujos neurônios (Figura 2) realizam uma operação de confluência de suas entradas:

$$v_j = \sum_{i=1}^n w_{ij} x_i \quad (1)$$

onde: v_j = entrada total recebida pelo neurônio;

w_{ij} = pesos sinápticos;

x_i = entradas do neurônio.

Uma função *threshold* (limiar) pode ser aplicada à soma ponderada, fornecendo uma maior variabilidade do grau de ativação deste neurônio. Esta função adiciona uma entrada -1 ao neurônio e pode ser incorporada a equação 1, adicionando-se $w_{j0}x_0$, que resulta em um somatório de $i=0$ a n :

$$v_j = \sum_{i=0}^n w_{ij} x_i \quad (2)$$

Aplicando-se uma função de ativação a esta equação (2), produz-se a saída deste neurônio:

$$y_j = f(v_j) \quad (3)$$

Geralmente a função de ativação é a função sigmóide:

$$f(v_j) = \frac{1}{1 + e^{-v_j}} \quad (4)$$

O aprendizado da rede consiste de um processo iterativo em que, dado um sinal de entrada, obtém-se uma resposta, ou estímulo, que é comparado a sua saída desejada (aprendizado supervisionado). Os pesos, que armazenam o conhecimento da rede, são reajustados, retropropagados para cada padrão apresentado, até um dado instante em que a função custo, ou erro (equação 5) atinja um mínimo global. Este algoritmo é denominado retropropagação padrão (*standard backpropagation*).

$$\varepsilon = \frac{1}{2} \sum_{j=1}^m (d_j - o_j)^2 \quad (5)$$

onde: d_j = saída desejada;

o_j = saída real;

m = número de neurônios na última camada.

Diversas variações deste algoritmo têm sido desenvolvidas, entre estas, *backpropagation with momentum* e RPROP. Este último algoritmo é uma variação do *standard backpropagation* que elimina a influência prejudicial da extensão da derivada parcial do erro em relação aos pesos. Como consequência, somente o sinal da derivada é considerado para indicar a direção de atualização dos pesos (Riedmiller, 1994). Para um maior detalhamento e compreensão das redes MLP e outras topologias em geral, ver Rumelhart (1989).

METODOLOGIA

A Tabela 1 apresenta um resumo dos parâmetros de caracterização da ETE. Os símbolos preenchidos indicam que estes parâmetros foram medidos diariamente no ponto de coleta especificado na primeira linha desta tabela. Estes parâmetros são relativos à operação da estação no período de 01/03/90 a 29/10/91.

Utilizou-se o simulador SNNS (Stuttgart Neural Network Simulator) (Stuttgart, 1995), desenvolvido no Instituto para Sistemas Paralelos e Distribuídos de Alta Performance da Universidade de Stuttgart para a definição das arquiteturas, aprendizado, validação e teste das redes testadas. Todos os parâmetros foram normalizados, sendo que os parâmetros não determinados, por quaisquer motivos, foram preenchidos, utilizando uma técnica descrita na próxima seção.

Tabela 1: Parâmetros determinados em seus respectivos pontos de coleta

parâmetros	entrada da estação	entrada decantador primário	entrada decantador secundário	saída da estação
Q (m³/dia)	●	-	-	-
Zn (mg/L)	●	-	-	-
pH	●	●	●	●
DBO (mg/L)	●	●	●	●
DQO (mg/L)	●	-	●	●
SS (mg/L)	●	●	●	●
SSV (mg/L)	●	●	●	●
sed. (mL/L)	●	●	●	●
cond. $\Omega^{-1}\text{cm}^{-1}$	●	●	●	●

Fonte: (Poch, 1993)

- parâmetros determinados.
- parâmetros não determinados no ponto de coleta.

Por serem medidos diariamente 29 parâmetros e tendo-se como objetivo a previsão destes mesmos parâmetros, baseados em informações passadas e presentes (parâmetros avaliados nos últimos dias), as redes possuem na camada de saída 29 neurônios. Foram definidas 5 topologias (arquiteturas) de redes, variando-se o número de dias utilizados para fazer a previsão e conseqüentemente, o número de nós de suas camadas de entrada. Assim, a rede com 29 nós na camada de entrada fez uso de somente 1 dia para a previsão dos parâmetros do dia seguinte da estação, e as redes com 58, 87, 116 e 145 nós utilizaram 2, 3, 4 e 5 dias, respectivamente, para a previsão destes parâmetros. Nesta fase do trabalho camadas intermediárias de neurônios não foram utilizadas.

A partir da definição do número de nós na camada de entrada e de neurônios na camada de saída para cada rede, foram determinados, aleatoriamente, três conjuntos de treinamento, validação e teste, sendo que 50% dos dados (padrões) foram reservados para a fase de treinamento, 25% para a fase de validação e 25% para a de teste.

Foram utilizados três algoritmos de treinamento diferentes, *standard backpropagation*, *backpropagation with momentum* e RPROP. A taxa de aprendizado utilizada foi variável, iniciando em 0,9 e reduzindo-a gradualmente até a taxa mínima de 0,01. Para o algoritmo *backpropagation with momentum*, utilizou-se a mesma taxa de aprendizado variável e o termo de momento foi fixado em 0,1. O número máximo de apresentações (ciclos) do conjunto de padrões de treinamento foi de 5000 e, e a cada 10 ciclos, apresentou-se o conjunto de validação. O conjunto de teste foi utilizado somente no final do treinamento.

RESULTADOS E DISCUSSÕES

A Figura 3 mostra a técnica utilizada para o preenchimento dos parâmetros ausentes, pois o simulador de redes neurais SNNS não permite intervalos de dados não preenchidos. O exemplo aqui ilustrado é um caso extremo, em que a vazão não foi medida por vários dias. Assim, uma função linear é traçada entre os dois extremos do intervalo ausente deste parâmetro. Este procedimento visou acompanhar a tendência dos dados, fato este não levado em consideração caso fosse adotado o procedimento usual, a média dos 4 últimos parâmetros.

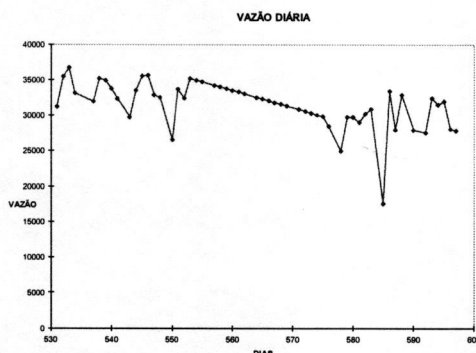


Figura 3: Preenchimento dos parâmetros ausentes.

As médias dos erros¹ de treinamento, validação e teste² para as redes treinadas com 3 conjuntos de padrões distintos estão apresentadas nas Figuras 4, 5 e 6, para cada um dos algoritmos de treinamento utilizados. O treinamento, validação e teste de cada arquitetura com 3 conjuntos de dados, extraídos aleatoriamente do conjunto total, visou eliminar erros causados por supostos intervalos não-representativos do comportamento global da estação.

Notou-se nestas figuras que, as médias dos erros quadráticos médios determinados por estas redes foram altos, embora isto não signifique um mau desempenho da rede. Muitas vezes estes erros podem estar associados a algumas saídas muito complexas (parâmetro de difícil modelagem) e ao grande número de neurônios na camada de saída, contribuindo sensivelmente para o aumento do somatório dos erros.

Observou-se também que, quanto mais nós de entrada (parâmetros para a previsão) foram acrescentados, menores os erros de treinamento e maiores os erros de validação e teste. Isto indicou que redes com mais nós na camada de entrada atingiram um alto grau de conhecimento sobre os padrões, mas não conseguiram generalizar suficientemente bem quando os conjuntos de validação e teste foram apresentados, ou seja, a rede decorou os

¹ Os erros referem-se ao Erro Quadrático Médio (MSE), dado por:

$$MSE = \frac{\left(\sum_{p=1}^m \sum_{i=1}^n (d_i - o_i)^2 \right)}{m}$$

onde: m = número de padrões de treinamento;
 n = número de neurônios na camada de saída;
 d_i = saída desejada no neurônio i;
 o_i = saída real do neurônio i.

² Os erros de treinamento, validação e teste apresentados correspondem aos ciclos cujos erros de validação foram os menores entre todos os ciclos da etapa de aprendizagem.

padrões apresentados, não tendo capacidade de generalizar novos padrões (o conjunto de teste). Assim, pode-se afirmar que a rede com 29 nós de entrada apresentou a melhor performance, visto que, apesar de não apresentar erros significativamente baixos, conseguiu generalizar melhor que as demais redes.

As Figuras 7, 8, 9, 10 e 11 apresentam uma comparação dos erros quadráticos médios para as 5 arquiteturas treinadas com os 3 algoritmos distintos. Observou-se que, para as redes com 58, 87, 116 e 145 nós de entrada (Figuras 8, 9, 10 e 11), o algoritmo RPROP mostrou-se mais eficiente para as fases de validação e teste à medida que nós são acrescentados na camada de entrada. Para a rede com 29 nós na entrada, os algoritmos *standard backpropagation* e *backpropagation with momentum* apresentaram uma ligeira vantagem na fase de teste.

É importante salientar que não foram utilizadas camadas intermediárias nas redes para que se pudesse verificar suas performances com arquiteturas as mais simples possíveis. Somente a partir dos erros determinados por estas redes, novas estruturas definidas.

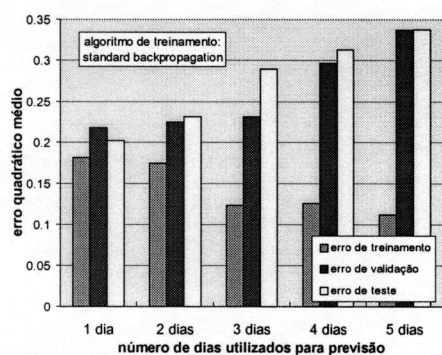


Figura 4: Erros das 5 arquiteturas utilizando algoritmo de treinamento *standard backpropagation*

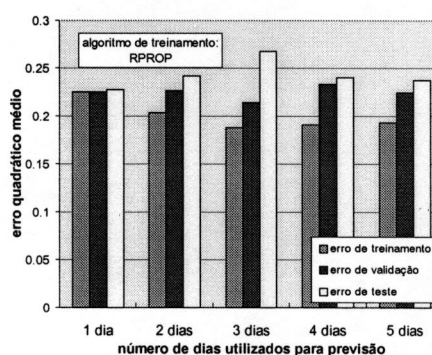


Figura 6: Erros das 5 arquiteturas utilizando algoritmo de treinamento RPROP

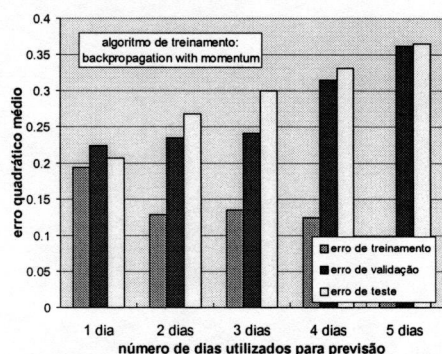


Figura 5: Erros das 5 arquiteturas utilizando algoritmo de treinamento *backpropagation with momentum*

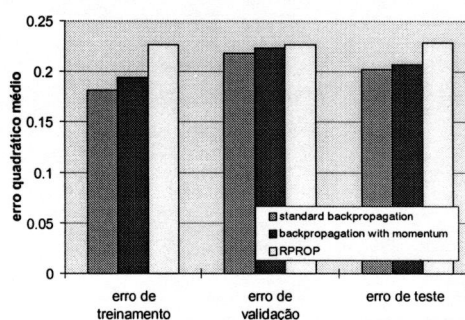


Figura 7: Erros para a rede com 29 nós na camada de entrada (1 dia para a previsão)

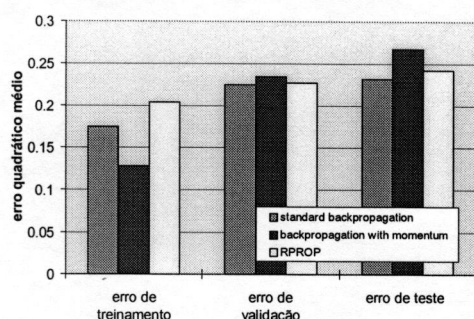


Figura 8: Erros para a rede com 58 nós na camada de entrada (2 dias para a previsão)

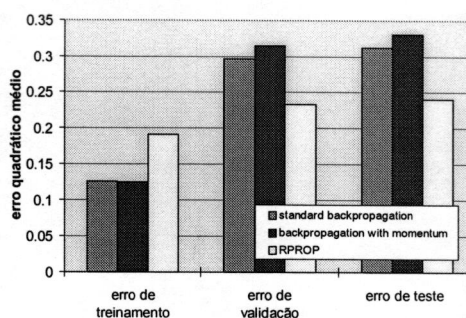


Figura 10: Erros para a rede com 116 nós na camada de entrada (4 dias para a previsão)

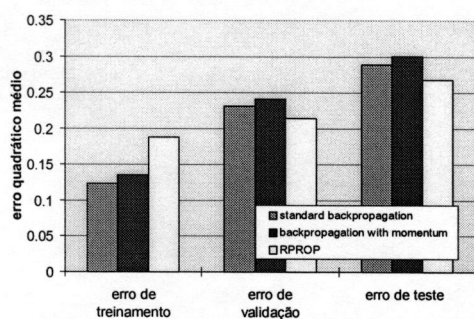


Figura 9: Erros para a rede com 87 nós na camada de entrada (3 dias para a previsão)

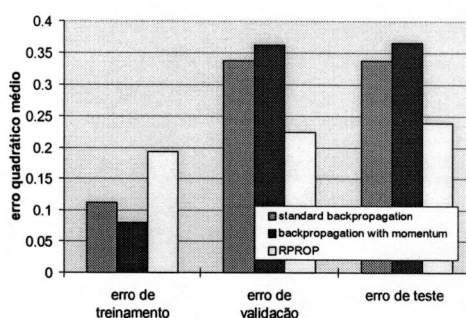


Figura 11: Erros para a rede com 145 nós na camada de entrada (5 dias para a previsão)

CONCLUSÕES E SUGESTÕES

O modelo de previsão proposto utilizando redes neurais mostrou-se bastante promissor na modelagem de sistemas biológicos de tratamento de esgotos, embora nenhum resultado realmente conclusivo possa ainda ser extraído. A próxima etapa deste estudo visará verificar a influência da adição de neurônios em camadas intermediárias. Segundo o teorema de Kolmogorov (Kolmogorov, 1963) não é necessário mais que três camadas para se aproximar qualquer tipo de função. Depois de terminada a etapa de previsão dos parâmetros, pretende-se viabilizar a emissão de diagnósticos da planta, diminuindo-se o número de nós na camada de entrada das redes, considerando-se somente os parâmetros, experimentais ou calculados, realmente relevantes à caracterização da planta.

AGRADECIMENTOS

Ao CNPq pelo concedimento da bolsa de doutorado e auxílio na realização deste trabalho.
Ao Prof. Ignasi R. Roda (Universidade Autônoma de Barcelona) pelas informações fornecidas sobre a ETE de Manresa, Espanha.

REFERÊNCIAS BIBLIOGRÁFICAS

1. BEALE, R.; JACKSON, T. **Neural computing: an introduction** IOP Publishing, Bristol, 1990.
2. BELANCHE, L. et al. A knowledge-based system for the diagnosis of waste-water treatment plant **Proceedings of the 5th international conference of industrial and engineering applications of AI and Experts Systems IEA/AIE**, Ed. Springer-Verlag, Alemanha, 1992.
3. BHAT, N. V. et al. Modelling chemical process systems via neural computation **IEEE Control Systems**, v. 10, n.3, p. 24-30, 1990.
4. CRUZ, A. J. G.; GIORDANO, R. C. **Relatório interno sobre redes neuronais** Universidade Federal de São Carlos, Centro de Ciências e Tecnologia, Programa de Pós-Graduação em Engenharia Química, São Carlos, 1995.
5. DINIZ, H. et al. **Stuttgart neural network simulator: manual de referência rápida** Instituto de Ciências Matemáticas de São Carlos - USP, São Carlos, 1996
6. HAYKIN, S. **Neural networks: a comprehensive foundation** Macmillan College Publishing Company, New York, 1994
7. KAWAI, H. et al. **Avaliação do desempenho de estações de tratamento de esgotos** CETESB, São Paulo, 38p, 1991.
8. KOLMOGOROV, A. N. On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. **Doklady Akademii Nauk SSSR**, v. 144, p. 679-681. (American Mathematical society Translation, 28: 55-59)
9. KROVVIDY, S.; WEE, W. E. Wastewater treatment systems from case-based reasoning. **Machine Learning**, v. 10, p. 341-363, 1993.
10. LAPOLLI et al. Utilização de técnicas de inteligência artificial e da teoria dos conjuntos difusos no controle operacional de estações de tratamento de esgotos. Simpósio Ítalo-brasileiro de Engenharia Sanitária e Ambiental, In: **Anais**. Rio de Janeiro, 1992.
11. MORRIS, A. J. et al. Artificial neural networks: studies in process modelling and control **Trans IchemE**, v. 72, p. 3-19, 1994.
12. POCH, M. **Faults in a urban waste water treatment plant** Banco de Dados do Departamento de Engenharia Química, Universidade autônoma de Barcelona, Espanha, 1993.
13. RIEDMILLER, M. Rprop **Relatório Técnico de Detalhes de Descrição e Implementação**, Universidade de Karlsruhe, janeiro, 1994.
14. RUMELHART et al. **Parallel distributed processing: explorations in the microstructure of cognition** v.1, MIT Press, 545p, 1989
15. SILVA JÚNIOR, R. J. C. A indústria da água como atividade econômica **Bio**, ABES, n. 1, jan/mar, p.22-33, 1996.
16. STUTTGART NEURAL NETWORK SIMULATOR **Implementation Manual** versão 4.1, Universidade de Stuttgart, Institute for Parallel and Distributed High Performance Systems (IPVR), 1995.
17. TINOCO FILHO Primeira página: o interesse municipal **Bio**, ABES, n. 1, jan/mar, p.1, 199.