

# IDENTIFICAÇÃO MULTIVARIÁVEL DE UMA COLUNA DE DESTILAÇÃO BASEADA EM REDE NEURAL E ALGORITMO *EXPECTATION-MAXIMIZATION*

LEANDRO DOS SANTOS COELHO

*Laboratório de Automação e Sistemas, Grupo Produtônica  
Pós-Graduação em Engenharia de Produção e Sistemas, PUCPR / PPGEPS  
Imaculada Conceição, 1155, CEP 80215-901 Curitiba, PR  
E-mails: leandro.coelho@pucpr.br*

EMERSON MARTIM E NEI HANSEN DE ALMEIDA

*Curso de Graduação em Engenharia Química  
Imaculada Conceição, 1155, CEP 80215-901 Curitiba, PR  
E-mails: emerson.martim@pucpr.br; nei.almeida@pucpr.br*

**Abstract** For complex systems, modeling using basic laws to determine their dynamic behavior is not always possible. An alternative to solve these problems is use concepts of systems identification. Through system identification it is possible to determine a mathematical model based on input and output data of the system. There are a wide variety of identification methods and model structures to choose for system nonlinear identification applications. Artificial neural networks have proved to be versatile and useful nonlinear models for industrial processes. This paper presents a design of neural network for identification of a distillation column. The proposed neural network is the radial basis function (NN-RBF) based on learning by expectation-maximization clustering algorithm. Numerical results indicate the potentialities of the NN-RBF for one-step-ahead forecasting of temperature outputs of an experimental distillation column.

**Keywords**—Radial basis function neural network, systems identification, expectation-maximization, distillation column.

**Resumo** Para sistemas complexos, a modelagem baseada em leis elementares para determinar o comportamento de sistemas dinâmicos nem sempre é possível. Uma alternativa para solucionar este problema é a utilização de conceitos de identificação de sistemas. Através da identificação de sistemas é possível determinar um modelo matemático baseado em dados experimentais das entradas e saídas do sistema. Existe uma grande variedade de métodos de identificação e estruturas de modelos para escolha em aplicação de identificação de sistemas não-lineares. As redes neurais artificiais têm provado serem versáteis e úteis como modelos não-lineares para processos industriais. Este artigo apresenta um projeto de rede neural para identificação de uma coluna de destilação. A rede neural proposta é a rede com função de base radial (RN-RBF) baseada em aprendizado por algoritmo de máximo das esperanças. Os resultados numéricos indicam as potencialidades da RN-RBF para previsão um passo a frente das saídas de temperatura de uma coluna de destilação experimental.

**Palavras-chave** Rede neural com função de base radial, identificação de sistemas, máximo das esperanças, coluna de destilação.

## 1 Introdução

A identificação de sistemas, seja ela realizada de forma *on-line* ou *off-line*, é uma parte essencial no projeto de sistemas de controle baseados em modelo. Em geral, a identificação de sistemas lineares é uma área bem estabelecida e através dos anos, muitas abordagens desenvolvidas já estão consolidadas na literatura (Foss *et al.*, 1998). Entretanto, em muitas situações práticas, a hipótese de linearidade não pode ser mantida devido à existência de não-linearidades ou elementos variantes no tempo. Em tais casos, é difícil aplicar os métodos quantitativos convencionais que descrevam adequadamente as características não-lineares e variantes no tempo de processos industriais.

Recentemente, o significativo desenvolvimento de metodologias da inteligência computacional, tais como redes neurais (Tan *et al.*, 2000), algoritmos evolutivos (Ahmad *et al.*, 1998) e sistemas nebulosos (Venkat e Gudi, 2002) têm providenciado

ferramentas para resolver problemas de identificação de sistemas dinâmicos com características complexas, tais como não-linearidades, atraso de transporte variante no tempo e dinâmica acoplada.

Uma abordagem relevante entre tantas outras para representação de sistemas complexos é a das redes neurais funções de base radial (RNs-RBF). As RNs-RBF são uma alternativa de modelo não-linear para problemas de regressão e classificação de padrões. Grande parte do interesse pelas RNs-RBF é proveniente do fato delas possuírem fundamentação estatística e são consideradas aproximadores universais de funções com dada precisão (Lo, 1998), realizando um mapeamento das entradas de rede de dimensão vetorial  $m$ , em uma saída de dimensão vetorial  $n$  de qualquer função vetorial contínua.

A literatura sobre a aplicação de RNs-RBF é vasta, a citar, por exemplo, as pesquisas de Tan *et al.* (1995), Abido e Abdel-Magid (1997), Yingwei *et al.* (1997), Billings e Hong (1998), Liu e Kadirkamanathan (1999), De Nicolao e Trecate (1999), Yu *et al.* (2000), Sheng *et al.* (2002) e Peng *et al.* (2003).

A contribuição deste trabalho é avaliar o desempenho de uma RN-RBF treinada com método de Maximização de Esperanças (*Expectation-Maximization*, EM). Neste contexto, a RN-RBF é avaliada em uma aplicação prática de identificação multivariável das variáveis de temperatura de uma coluna de destilação em configuração *SIMO* (*Single Input Multiple Output*). A estrutura de modelo escolhida foi o modelo multivariável *NARX* (*Nonlinear AutoRegressive with eXogenous inputs*).

O artigo é organizado da seguinte forma. Os fundamentos das RNs-RBF e o método EM são descritos na seção 2. A descrição da coluna de destilação experimental e a análise dos resultados obtidos da aplicação da RN-RBF são mencionadas respectivamente nas seções 3 e 4. A conclusão é apresentada na seção 5.

## 2 Fundamentos das RNs-RBF

As redes neurais artificiais são motivadas por sistemas neurais biológicos, com intenção de simular, mesmo que rudimentarmente, a forma como o cérebro aprende, recorda e processa informações. As redes neurais artificiais são projetadas como um sistema interconectado de elementos de processamento (neurônios), cada um dos quais com um número limitado de entradas e saídas.

O processamento da entrada de uma rede neural se realiza de acordo com sua topologia e o tipo de ativação dos neurônios. Em uma RN-RBF clássica, este aspecto consiste na alimentação adiante (*feedforward network*) de uma entrada de dimensão  $n$  em uma camada oculta (intermediária), que possui  $M$  neurônios ocultos que representam uma função de ativação baseada em uma medida de similaridade entre os vetores de entrada e os protótipos. Cada neurônio da camada oculta possui um protótipo. As saídas da camada oculta são interligadas às unidades da camada de saída por ramos com pesos sinápticos  $w_j$ ,  $j=1, \dots, M$ , que ponderam cada saída dos neurônios presentes na camada oculta. Para cada unidade da camada de saída é calculada uma soma ponderada das saídas da camada oculta. A resposta das camadas de saída forma os elementos do vetor de resposta de uma RN-RBF.

Uma função de base radial,  $f$ , possui saída simétrica em torno de um centro associado,  $\mathbf{m}_c$ , tal que  $f_c = f(\|x - \mathbf{m}_c\|; \mathbf{s})$ , onde  $\|\cdot\|$  é um vetor norma. Muitas funções são utilizadas como funções de base, na camada intermediária de uma RN-RBF, tais como funções de aproximação cúbica, multiquadrática e *thin plate spline*.

Uma condição suficiente para uma função ser candidata à função de base radial é a de ser radialmente simétrica e a sua primeira derivada deve ser monotônica (Poggio e Girosi, 1990). As funções Gaussianas, uma das funções de base mais utilizadas

na literatura, são caracterizadas por um parâmetro de escala (ou comprimento),  $\mathbf{s}$ , e podem ser representadas por

$$f_c(x) = f(\|x - \mathbf{m}_c\|; \mathbf{s}) \quad (1)$$

Um conjunto de funções de base radial pode servir como uma base para a representação de uma vasta classe de funções que são expressas como combinações lineares de funções de base radial escolhidas, tal que:

$$y(x) = \sum_{j=1}^M \mathbf{w}_j \cdot f(\|x - \mathbf{m}_c\|; \mathbf{s}) \quad (2)$$

Uma RN-RBF, conforme apresentada na Fig. 1, é um agrupamento de equações na forma de uma rede neural *feedforward* com três camadas: a(s) entrada(s), camada oculta (intermediária) e camada(s) de saída. Cada unidade oculta representa uma função de base radial simples com comprimento e posição do centro associados. As unidades ocultas são algumas vezes denominadas de centróides ou núcleos. Cada unidade de saída executa uma soma ponderada de unidades ocultas, utilizando os valores de  $\mathbf{w}_j$  como pesos sinápticos.

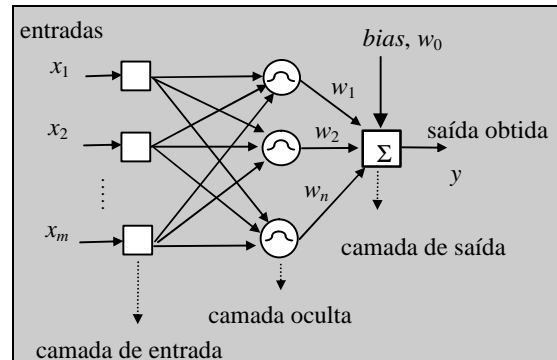


Fig. 1. RN-RBF com uma camada intermediária.

A abordagem de encontrar os componentes individuais sendo dada uma superposição de Gaussianas foi explorado nos anos 1960. Subseqüentemente, a utilização de funções de base radial para interpolação numérica e aproximação de funções foi demonstrada por Powell (1985). O projeto da RN-RBF adotado, neste artigo, é o do treinamento em dois estágios, que são os seguintes: (i) determinar o número de centros e dos valores de  $\mathbf{m}_c$  e  $\mathbf{s}_j$ , e (ii) determinar os pesos das unidades da saída para os centros e comprimentos obtidos no estágio (i).

No primeiro estágio pode-se utilizar um algoritmo de aprendizado supervisionado ou não-supervisionado. Neste artigo, adota-se o método de agrupamento baseado em EM Gaussiano (Bishop, 1995) para o ajuste dos centros das funções de ativação Gaussianas. No segundo estágio, o ajuste

dos pesos das unidades de saída é realizado pela técnica da pseudo-inversa (Golub e Van Loan, 1983). A seguir são apresentados os fundamentos do algoritmo EM.

### 2.1. Fundamentos do algoritmo EM

O algoritmo EM é um procedimento iterativo utilizado para o cálculo de estimativas de Máxima Verossimilhança (*Maximum Likelihood*, ML) em situações em que as observações possam ser consideradas como tendo dados incompletos.

O algoritmo EM surgiu da unificação de uma série de trabalhos aparentemente sem relação e apresentado com esse nome por Demspster *et al.* (1977). Depois disto, este algoritmo permaneceu por muito tempo esquecido. Basicamente, se alguma variável foi algumas vezes observada e outras não, pode-se utilizar os casos para os quais foi observada para aprender a prever seus valores quando não. O algoritmo EM realiza esta tarefa, mas, também pode ser utilizado para variáveis cujos valores nunca foram observados, sempre e quando seja conhecida a forma geral da distribuição de probabilidade das variáveis (Luna, 2004).

Em geral os parâmetros descrevem as características de uma população. Seus valores são estimados de amostras coletadas dessa população. O algoritmo EM faz uma estimativa de máxima verossimilhança dos parâmetros, ou seja, estima parâmetros que sejam os mais consistentes como os dados da amostra no sentido de maximizar a função de verossimilhança (Luna, 2004).

Antes de introduzir o algoritmo EM, define-se função de verossimilhança e estimador de máxima verossimilhança, conforme Luna (2004) para o caso geral de  $n$  variáveis aleatórias que dependem do parâmetro  $\mathbf{q}$ , que pode até ser um vetor de parâmetros.

**Definição 1** (Função de Verossimilhança): A função de verossimilhança de  $n$  variáveis aleatórias  $X_1, X_2, \dots, X_n$  é definida como a densidade conjunta de  $n$  variáveis aleatórias, digamos  $f_{X_1, \dots, X_n}(x_1, x_2, \dots, x_n; \mathbf{q})$  considerada como uma função de  $\mathbf{q}$ . Em particular, se  $X_1, X_2, \dots, X_n$  é uma amostra aleatória da densidade  $f(x; \mathbf{q})$ , então a função de verossimilhança é  $f(x_1; \mathbf{q}), f(x_2; \mathbf{q}), \dots, f(x_n; \mathbf{q})$ . A notação utilizada, neste trabalho, para a função de verossimilhança é  $L(\mathbf{q}; x_1, x_2, \dots, x_n)$ .

**Definição 2** (Estimador de Máxima Verossimilhança): Seja  $L(\mathbf{q}; x_1, x_2, \dots, x_n)$  a função de verossimilhança para as variáveis aleatórias  $X_1, X_2, \dots, X_n$ . Se  $\hat{\mathbf{q}}$  é valor de  $\mathbf{q}$  em  $\bar{\Theta}$  (universo dos  $\mathbf{q}$ ) que maximiza  $L(\mathbf{q})$ , então  $\hat{\Theta} = \mathbf{u}(X_1, X_2, \dots, X_n)$  é o estimador de máxima verossimilhança de  $\mathbf{q}$ . Neste caso,  $\hat{\mathbf{q}} = \mathbf{u}(x_1, x_2, \dots, x_n)$  é a máxima verossimilhança de  $\mathbf{q}$  para a amostra  $x_1, x_2, \dots, x_n$ .

O algoritmo de EM pode ser resumido em dois passos, que são os seguintes (Luna, 2004):

**Passo E (Expectation):** encontra-se os valores esperados das estatísticas suficientes para os dados completos  $Y$ , dado os dados incompletos  $Z$  e as estimativas atuais dos parâmetros do modelo.

**Passo M (Maximization):** a função de máxima verossimilhança é maximizada assumindo que os dados incompletos  $Z$  são conhecidos.

O algoritmo EM pode ser utilizado como um eficiente método de agrupamento (*clustering*) de dados. O EM na concepção Gaussiana particiona o conjunto de dados em grupos pela determinação do modelo de mistura (*mixture models*) de Gaussianas do conjunto de dados (dados de estimação) de entrada da RN-RBF.

Cada Gaussiana possui a média e a matriz de covariância. A função objetivo  $J_{EM}$  que o EM otimiza é definida por Hamerly e Elkan (2002) e é dada por:

$$J_{EM} = - \sum_{p=1}^{N_p} \log \left( \sum_{k=1}^K p(z_p | m_k) p(m_k) \right) \quad (3)$$

onde  $p(z_p | m_k)$  é a probabilidade de  $z_p$ , dado que a probabilidade é gerada por uma distribuição Gaussiana com centróide  $m_k$ , e  $p(m_k)$  é a probabilidade a priori do centróide  $m_k$ .

As funções de pertinência e ponderação para o algoritmo EM são definidas por (Hamerly e Elkan, 2002):

$$u(m_k | z_p) = \frac{p(z_p | m_k) p(m_k)}{p(z_p)} \quad (4)$$

$$w(z_p) = 1 \quad (5)$$

Uma vez que o EM tem função de pertinência e uma função de ponderação constante. O algoritmo EM começa com uma estimativa inicial dos parâmetros. Então, um *Passo E* é aplicado, onde os valores dos dados conhecidos são utilizados para calcular os valores esperados dos dados desconhecidos (Hamerly, 2003). Isto é seguido pelo *Passo M*, onde os valores conhecidos e esperados dos dados são utilizados para gerar uma nova estimativa dos dados. Os *Passos E* e *M* são repetidos até que um critério de convergência seja atingido.

Neste trabalho, o objetivo do uso do algoritmo EM é obter os valores para os centros das funções de ativação Gaussianas da RN-RBF. Detalhes sobre a formulação matemática do procedimento iterativo do algoritmo EM são apresentados em Bishop (1995), McLachlan e Krishnan (1997), Borman (2004) e Luna (2004).

### 3 Descrição da Coluna de Destilação

A operação a vácuo de uma coluna de destilação com recheio estruturado permite então realizar não apenas uma simples destilação de óleos essenciais, mas também uma maior purificação dos produtos (Holland, 1981). Além de ser mais eficiente do ponto de vista do produto, a operação a vácuo de uma coluna de destilação com recheio estruturado, otimiza a quantidade de insumos necessários para a produção do óleo essencial. O fato de produzir eficientemente o óleo (menor quantidade de insumos) vai de encontro aos objetivos globais do referido projeto, consome menos insumos e gera-se menor quantidade de subprodutos (Almeida, 1994).

Na Fig. 2 é apresentado um diagrama do processo de destilação a vácuo com recheio estruturado para a planta piloto abordada neste trabalho.

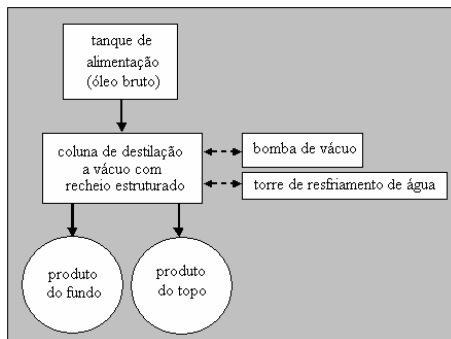
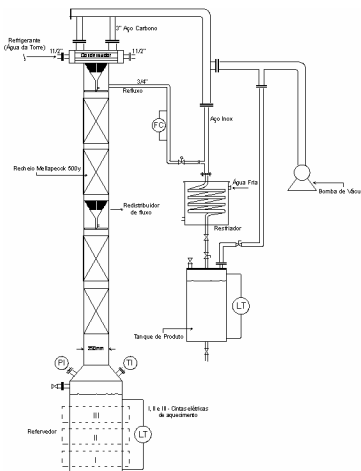


Fig. 2. Diagrama esquemático do processo de purificação (coluna de destilação a vácuo com recheio estruturado).

Na Fig. 3 são apresentados o *layout* do processo de destilação a vácuo com recheio estruturado para a planta piloto (localizada na Fazenda Experimental Galha Azul da PUCPR) e sua fotografia, que foi alvo deste trabalho. Neste contexto, o objetivo deste trabalho é a elaboração de um modelo matemático baseado em RN-RBF sobre o processo de destilação a vácuo com recheio estruturado orientado para a destilação da terebentina.



(a) *layout* do processo de fabricação.



(b) fotografia da coluna de destilação.

Fig. 3. Coluna de destilação a vácuo com recheio estruturado da planta piloto da Fazenda Experimental Galha Azul da PUCPR.

### 4 Resultados da Identificação usando RN-RBF

No procedimento de identificação da coluna de destilação foi utilizada uma configuração *SIMO* (*Single Input Multiple Output*) que consiste de 1 entrada (volume de terebentina em litros),  $u(t)$ , e 3 saídas ( $y_1(t)$ : temperatura do refeedor,  $y_2(t)$ : temperatura do ponto quente, e  $y_3(t)$ : temperatura do ponto frio). Os dados coletados experimentalmente foram com um período de amostragem de aproximadamente 15 minutos.

Para o procedimento de identificação foi escolhido um modelo matemático para a representação do processo. A estrutura de modelo escolhida foi o modelo multivariável *NARX* (*Nonlinear AutoRegressive with eXogenous inputs*), este representado na Fig. 4.

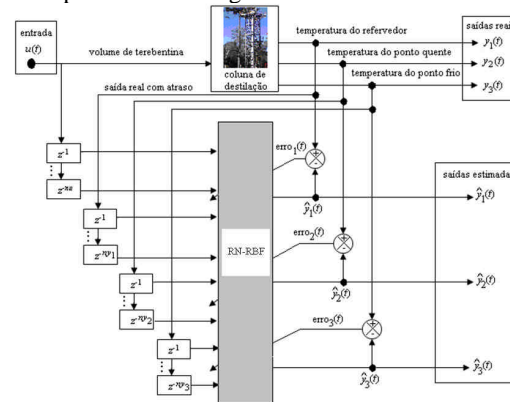


Fig. 4. Configuração de identificação série-paralelo usando uma RN-RBF com treinamento por algoritmo EM.

Nas simulações para identificação usando a RN-RBF, foram utilizadas as amostras 1 a 58 na fase de estimação (etapa de treinamento) e as amostras 59 a 116 na fase de validação (etapa de testes). Os vetores de entrada para a RN-RBF foram selecionados tal que o modelo obtido fosse com entradas de primeira ordem  $nu=1$  e  $ny_1=ny_2=ny_3=1$ . Os dados da entrada utilizados na identificação são apresentados na Fig. 5. Nota-se que houveram retiradas de volume da coluna de destilação num total de 8 frações, em

determinadas temperaturas, conforme apresentado na Tabela 1.

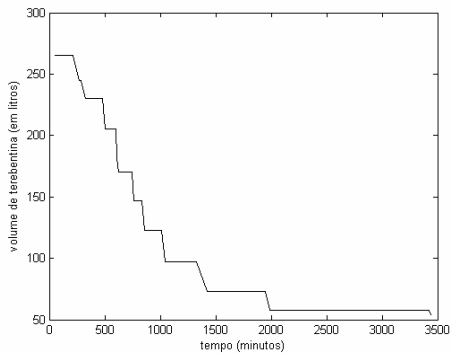


Fig. 5. Dados de entrada utilizados na identificação pela RN-RBF.

Tabela 1. Temperatura de ebulição dos componentes da terebentina.

Frações da terebentina	Temperatura (em °C) à 760mmHg
Alfa pineno	155 - 156
Canfeno	158,5 - 159,5
Beta pineno	165 - 166
Mirceno	167
Limoneno	175,5 - 176
Felandreno	178 - 179
7metil chavicol	216
8-alfa terpineol	218,8 - 219,4

O resultado obtido na identificação pela RN-RBF foi melhor quando 5 funções Gaussianas foram utilizadas para cada entrada das RN-RBF, conforme observado na Tabela 2. Este resultado é apresentado na Fig. 6.

Tabela 2. Resultados da identificação da coluna de destilação usando RN-RBF.

etapa de estimação			
Gaussianas	$MSE_1$	$MSE_2$	$MSE_3$
2	98,4746	32,7734	24,2920
3	54,5129	12,8125	9,1811
4	39,9773	11,3057	7,5304
<b>5</b>	<b>33,8112</b>	<b>11,3131</b>	<b>7,5297</b>
6	35,9215	11,3237	7,6799
7	45,5809	13,6665	9,3564
etapa de validação			
Gaussianas	$MSE_1$	$MSE_2$	$MSE_3$
2	100,3028	27,1014	20,1220
3	37,2253	7,1124	6,1619
4	33,5281	7,1772	4,6968
<b>5</b>	<b>32,6031</b>	<b>7,1382</b>	<b>4,6966</b>
6	34,0121	7,2608	4,7896
7	50,2173	9,0234	6,5352
convenções			
Gaussianas: número de funções de ativação Gaussianas para cada entrada da RBF.			
$MSE_i$ (Mean Square Error): erro médio quadrático para a saída $i$ da coluna de destilação.			

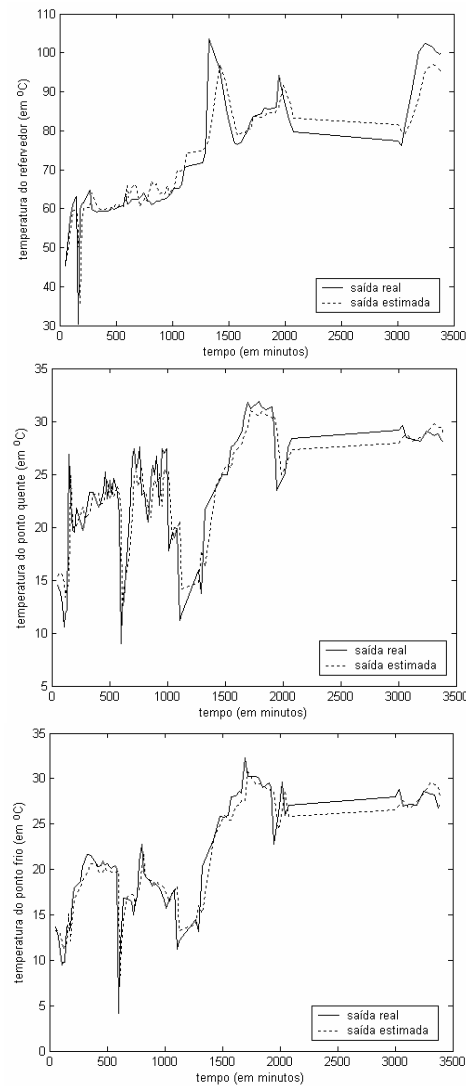


Fig. 6. Resultados para identificação usando 5 funções Gaussianas para cada entrada da RN-RBF.

## 5 Conclusão

Neste trabalho foram apresentados os fundamentos e um procedimento de identificação multivariável usando uma RN-RBF. Os resultados deste trabalho mostram que a RN-RBF aproxima o comportamento não-linear dos sinais de saída da coluna de destilação em uma configuração SIMO, principalmente quando o projeto da RN-RBF utiliza mais que 4 funções Gaussianas na camada intermediária.

Em futura pesquisa, a abordagem de RN-RBF será aprimorada para previsões  $k$ -passos à frente para a coluna de destilação usando método de otimização para obtenção de um melhor ajuste do parâmetro de escala,  $\mathbf{s}$ , das funções Gaussianas da RN-RBF.

## Referências Bibliográficas

Abido, M. A. e Abdel-Magid, Y. L. (1997). On-line identification of synchronous machines using

- radial basis function neural networks, *IEEE Transactions on Power Systems*, **12**(4): 1500-1506.
- Ahmad, M., Zhang, L. e Readle, J. C. (1998). Parameter estimation of a heating system using a genetic algorithm, *UKACC International Conference on Control*, Swansea, UK, pp. 195-200.
- Almeida, N. H. (1994). Determinação de eficiência do recheio estruturado mellapak 500 Y na destilação da terebintina em coluna piloto, *Tese de doutorado*, Universidade de São Paulo, SP.
- Billings, S. A. e Hong, X. (1998). Dual-orthogonal radial basis function networks for nonlinear time series, *Neural Networks*, **11**(3): 479-493.
- Bishop, C. M. (1995). *Neural networks for pattern recognition*, Oxford University Press, Oxford, UK.
- Borman, S. (2004). The expectation maximization algorithm: a short tutorial, *Technical Report*, Laboratory for Image and Signal Analysis, Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN, USA.
- De Nicolao, G. e Trecate, G. F. (1999). Consistent identification of NARX models via regularization networks, *IEEE Transactions on Automatic Control*, **44**(11): 2045-2049.
- Dempster, A. (1968). A generalization of Bayesian inference, *Journal of the Royal Statistical Society (Series B)*, **30**: 205-267.
- Foss, B. A., Lohmann, B. e Marquardt, W. (1998). A field study of the industrial modeling process, *Journal of Process Control*, **8**(5-6): 325-338.
- Golub, G. H. e Van Loan, C. F. (1983). *Matrix computations*, Baltimore: The John Hopkins University Press.
- Hamerly, G. e Elkan, C. (2002). Alternative to the k-means algorithm that find better clusterings, *Proceedings of the ACM Conference on Information and Knowledge Management*, McLean, VA, USA, pp. 600-607.
- Hammerly, G. (2003). Learning structure and concepts in data using data clustering, *PhD thesis*, University of California, San Diego, USA.
- Holland, C. D. (1981). *Fundamentals of multicomponent distillation*, McGraw Hill, Chemical Engineering Series.
- Liu, G. P. e Kadiramanathan, V. (1999). Multiobjective criteria for neural network structure selection and identification of nonlinear systems using genetic algorithms, *IEE Proceedings-Control Theory and Applications*, **146**(5): 373-382.
- Lo, J. T. -H. (1998). Multilayer perceptrons and radial basis functions are universal robust approximators, *International Joint Conference on Neural Networks*, Anchorage, AK, USA, v. 2, pp. 1311-1314.
- Luna, J. E. O. (2004). Algoritmos EM para aprendizagem de redes Bayesianas a partir de dados incompletos, *Dissertação de mestrado*, Mestrado em Ciência da Computação, UFMS, Campo Grande, MS.
- McLachlan, G. e Krishnan, T. (1997). *The EM algorithm and extensions*. Wiley series in probability and statistics, JohnWiley & Sons.
- Peng, H., Ozaki, T., Ozaki, V. H. e Toyoda, Y. (2003). A parameter optimization method for radial basis function type models, *IEEE Transactions on Neural Networks*, **14**(2): 432-438.
- Poggio, T. e Girosi, F. (1990). Networks for approximation and learning, *Proceedings of the IEEE*, **78**(9): 1481-1497.
- Powell, M. J. D. (1985). Radial basis functions for multivariable interpolation: a review, *IMA Conference on Algorithms for the Approximation of Functions and Data*, Shrivvenham, UK, pp. 143-167.
- Rendner, R. e Walker, H. (1984). Mixture densities, maximum likelihood and the EM algorithm, *SIAM Review*, **26**(2): 195-239.
- Shen, C., Cao, G. -Y. e Zhu, X. -J. (2002). Nonlinear modeling of MCFC stack based on RBF neural networks identification, *Simulation Modelling Practice and Theory*, **10**(1-2): 109-119.
- Tan, S., Hao, J. e Vanderwalle, J. (1995). Efficient identification of RBF neural net models for nonlinear discrete-time multivariable dynamical systems, *Neurocomputing*, **9**(1): 11-26.
- Tan, Y., Dang, X., Liang, F. e Su, S. -Y. (2000). Dynamic wavelet neural network for nonlinear dynamic system identification, *Proceedings of the IEEE International Conference on Control Applications*, Anchorage, AL, USA, pp. 214-219.
- Venkat, A. N. e Gudi, R. D. (2002). Fuzzy segregation-based identification and control of nonlinear dynamic systems, *Ind. Eng. Chem. Res.*, **41**: 538-552.
- Yingwei, L., Sundararajan, N. e Saratchandran, P. (1997). Identification of time-varying nonlinear systems using minimal radial basis function neural networks, *IEE Proc-Control Theory Appl.*, **144**(2): 202-208.
- Yu, D. L., Gomm, J. B. e Williams, D. (2000). Neural model input selection for a MIMO chemical process, *Engineering Applications of Artificial Intelligence*, **13**(1): 15-23.