

Identificação de Processo Real Não Linear do Tipo Hammerstein utilizando Redes Neurais com Dinâmica Externa e com Dinâmica Interna

R. C. L. de Oliveira¹, R.P.O Soares¹, C.E.U.Costa¹

¹Lab. de Controle e Sistemas

Depart. de Engenharia Elétrica, Univ. Federal do Pará

Caixa Postal 8619, 66.075-970

Belém, Pará, Brasil

E-mails: limao@ufpa.com.br, rosana@nautilus.com.br, ueno@amazon.com.br

Abstract

Neural Networks have been used to identify nonlinear dynamic systems because they are genuine nonlinear black box models and they have ability to approximate complicated nonlinear relations. This paper presents two neural net models, one with external dynamic, represented by a regression vector and the other with internal dynamic represented by dynamic neurons which are used to identify nonlinear systems. They has a feedforward topology and the learning algorithm used is backpropagation. Results attained in the identification of a real block-oriented nonlinear dynamic process are used to evaluate characteristic of each these neural net models.

retropropagação do erro (*backpropagation*) desenvolvido por Rumelhart, onde os pesos e bias da RNA são ajustados com o objetivo de minimizar o erro médio quadrático da rede.

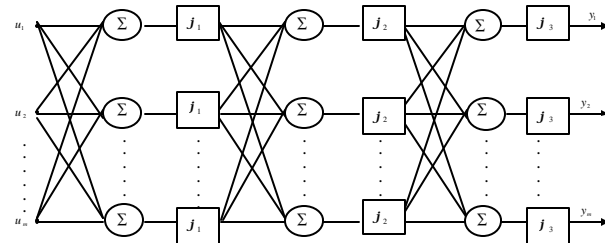


Figura 1: Rede Neural Direta Multicamadas

1. Introdução

O problema de identificação consiste em desenvolver um modelo da planta a partir de registros operacionais, ajustando seus parâmetros de forma que uma função custo seja otimizada, baseando-se no erro entre as saídas da planta e do modelo. Desde que as funções não lineares usadas para representar a planta pertençam a uma classe conhecida no domínio de interesse, a estrutura do modelo de identificação pode ser similar àquela da planta. A complexidade e diversidade dos sistemas não lineares dificulta a utilização dos métodos de identificação de forma genérica, contudo para uma classe de sistemas orientados a blocos, as Redes Neurais Artificiais (RNA's) têm se mostrado uma excelente ferramenta. Redes Diretas Multicamadas (figura 1), rapidamente tornaram-se um dos tipos mais utilizados de redes neurais, principalmente na área de sistemas e controle [2][3][4].

Para a maioria dos problemas que podem ser solucionados usando este tipo de RNA, a arquitetura é composta por um conjunto de padrões de entrada, uma ou duas camadas intermediárias e uma camada de saída e o algoritmo de treinamento utilizado é o de

2. Identificação Neural

A estrutura utilizada para treinar a RNA de modo que a mesma represente a dinâmica direta da planta, é mostrada na figura 2. A RNA é colocada em paralelo com a planta e o erro entre a saída do sistema e a saída da rede é utilizado como sinal de ajuste dos pesos da RNA. Esta arquitetura caracteriza um aprendizado supervisionado, onde a planta fornece os valores desejados (suas saídas) que servem para adaptar a RNA. No caso particular de uma RNA direta a retropropagação do erro de predição (*backpropagation*) através da rede fornece um possível algoritmo de aprendizado.

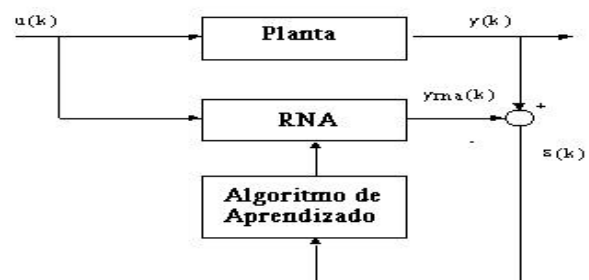


Figura 2: Identificação Neural com Modelagem Direta

A representação da dinâmica na RNA pode ser feita através do conjunto de sinais atrasados no tempo, na entrada da rede (vetor regressor), sendo também possível utilizar a representação através de neurônios dinâmicos, com a passagem do modelo entrada - saída para o modelo no espaço de estados. A figura 3 mostra um modelo geral do neurônio artificial [6]. Este modelo também possibilita a inclusão de dinâmica através da função de ativação, responsável pelo estado do neurônio, e apresenta a função de saída estática que gera o sinal de saída do neurônio.

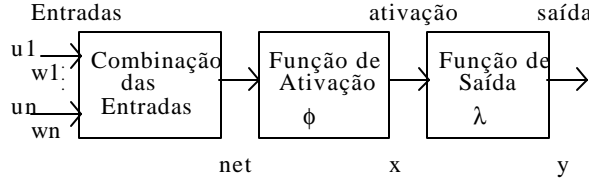


Figura 3 : Modelo de neurônio artificial

3. Sistemas Orientados a Blocos

Existem duas linhas mestras no desenvolvimento de métodos para identificar processos não-lineares. A primeira delas é o uso de uma descrição geral para representar o processo não linear. A segunda linha é considerar classes especiais de modelos não-lineares. Uma das classes de sistemas não lineares mais estudadas é a de sistemas orientados a blocos onde se considera modelos não lineares formados pela combinação de blocos não lineares estáticos e blocos dinâmicos lineares [2]. Exemplos destes modelos são : modelo de Wiener, modelo de Hammerstein, modelo Wiener-Hammerstein e modelo Hammerstein-Wiener [5], como observados na figura 4. Nestes modelos apenas os sinais de entrada e de saída são disponíveis. Estes modelos são considerados modelos tipo caixa-cinza porque existe o conhecimento prévio da sua estrutura.

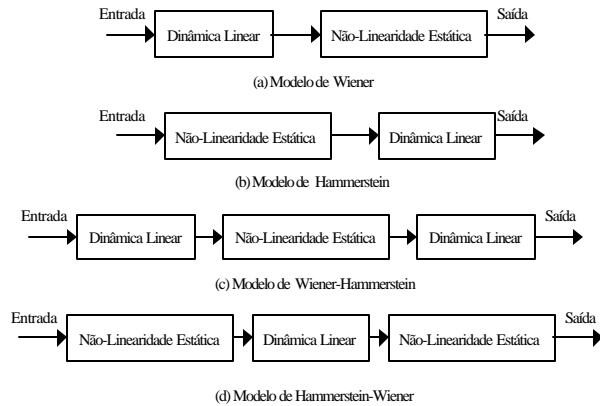


Figura 4 : Modelos Não Lineares orientados a blocos

Na literatura, as RNA's estão sendo bastante exploradas, como ferramenta para a identificação de processos não lineares que apresentam estas características [3]

4. RNA com Dinâmica Externa

Dado um processo, representado por,

$$y(k) = f[y(k-1), \dots, y(k-n), u(k-1), \dots, u(k-m)] \quad (1)$$

onde k é o instante de amostragem, $u(k)$ é o sinal de entrada, $y(k)$ é o sinal de saída, e $f(\cdot)$ é a função não linear de saída que realiza o mapeamento $\Re^{m+n} \rightarrow \Re^1$, o mesmo pode ter um modelo representado pela RNA da figura 5.

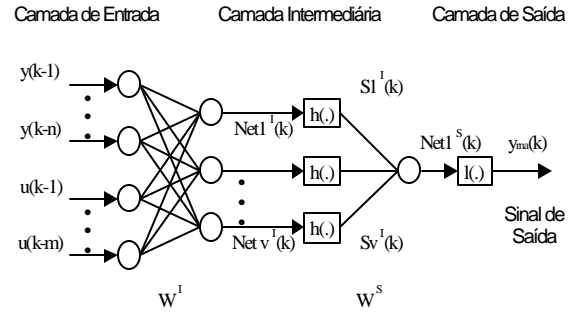


Figura 5 : RNA com Dinâmica Externa

Da figura 5 tem-se que,

$$S_j(k) = h[net_j(k)] \quad (2)$$

$$net_j(k) = w_{j1}^I u_{rna}(k) \quad (3)$$

onde,

$$u_{rna}(k) = [y(k-1), \dots, y(k-n), u(k-1), \dots, u(k-m)] \quad (4)$$

O sinal de saída da RNA será,

$$y(k) = l[W^S h[W^I u_{rna}(k-1)]] \quad (5)$$

$$W^S = \begin{bmatrix} w_{11}^S \\ w_{21}^S \\ \vdots \\ w_{v1}^S \end{bmatrix}; \quad W^I = \begin{bmatrix} w_{11}^I & w_{12}^I & \dots & w_{1v}^I \\ w_{21}^I & w_{22}^I & \dots & w_{2v}^I \\ \vdots & \vdots & \dots & \vdots \\ w_{(n+m)1}^I & w_{(n+m)2}^I & \dots & w_{(n+m)v}^I \end{bmatrix} \quad (6)$$

ou,

$$y(k) = N[W, y(k-1), u(k-1)] \quad (7)$$

que formaliza o modelo dado pelas equação (1). O algoritmo de aprendizado ajustará os elementos de W de tal forma que $N(\cdot) \approx f(\cdot)$.

5. RNA com Dinâmica Interna

Dado um processo representado por,

$$x(k+1) = Ax(k) + Bu(k) \quad (8)$$

$$y(k) = \Phi(x(k)) \quad (9)$$

onde k é o instante de amostragem, A é a matriz característica $n \times n$, B é o vetor entrada $n \times 1$, $u(k)$ é o sinal de entrada, $x(k)$ é o vetor de estados $n \times 1$, $y(k)$ é o sinal de saída, e $\Phi(\cdot)$ é a função não-linear de saída que realiza o mapeamento $\mathcal{R}^n \rightarrow \mathcal{R}^1$, o mesmo pode ter um modelo representado pela RNA da figura 6.

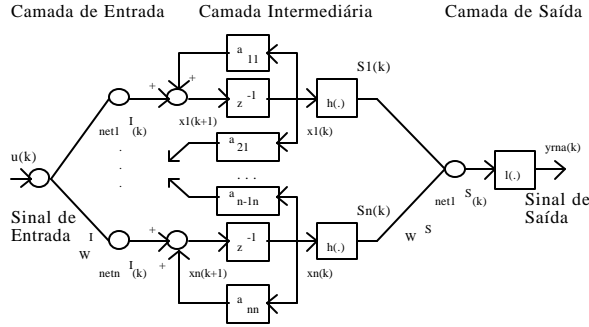


Figura 6 : RNA com Dinâmica Interna

Da figura 6 tem-se que,

$$S_i(k) = h[x_i(k)] \quad (10)$$

$$x_i(k+1) = a_{i1}x_1(k) + \dots + a_{ii}x_i(k) + \dots + a_{in}x_n(k) + net_i(k) \quad (11)$$

$$net_i(k) = w_{i1}^I u(k) \quad (12)$$

ou,

$$x_i(k+1) = a_{i1}x_1(k) + \dots + a_{in}x_n(k) + w_{i1}^I u(k) \quad (13)$$

O sinal de saída da RNA será,

$$y(k) = l[W^S h[Ax(k-1) + W^I u(k-1)]] \quad (14)$$

com,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}; \quad (15)$$

$$W^S = \begin{bmatrix} w_{11}^S \\ w_{21}^S \\ \vdots \\ w_{n1}^S \end{bmatrix}; \quad W^I = [w_{11}^I \quad w_{12}^I \quad \dots \quad w_{1m}^I] \quad (16)$$

ou,

$$y(k) = N[A, W, x(k-1), u(k-1)] \quad (17)$$

formalizando o modelo dado pelas equações (8) e (9). O algoritmo de aprendizado (De Oliveira *et alii*, 1997) ajustará os elementos de W e A de forma que $N(\cdot) \approx \Phi(\cdot)$.

6. Resultados

Para avaliar o desempenho de RNA's com Dinâmica Externa e Dinâmica Interna, na tarefa de identificar

processos não lineares da classe de sistemas orientados a blocos, foram realizados experimentos em um processo disponível no Laboratório de Controle e Sistemas - LACOS - do Departamento de Engenharia Elétrica da Universidade Federal do Pará.

O processo real utilizado, trata-se do equipamento PCS327 da Feedback Instruments Ltd., que implementa através de amplificadores operacionais várias funções de transferências. O equipamento é mostrado na figura 7.

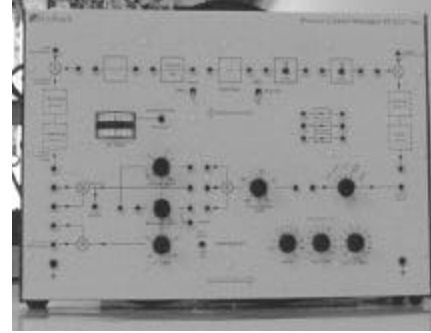


Figura 7: Equipamento PCS327 da Feedback Instruments Ltd.

A planta configurada no equipamento foi do tipo mostrado na figura abaixo

Modelo do Fabricante

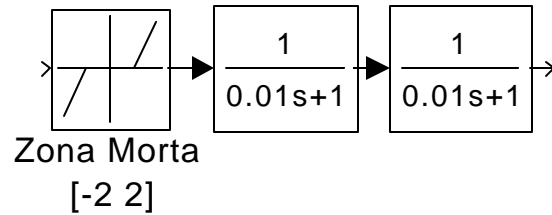


Figura 8 : Diagrama de Blocos da planta

Os padrões de treinamento das redes, foram obtidos da planta real. Na aquisição desses dados foi utilizado um intervalo de amostragem de 0.02 segundos e obtidas 301 amostras as quais foram utilizadas para treinamento. Os padrões de treinamento foram normalizados e o critério de parada utilizado foi a somatória do erro quadrático igual a 1,0.

A arquitetura da rede com dinâmica externa envolveu um vetor regressor do tipo $[y(k) \ y(k-1) \ u(k)]$, uma camada intermediária com 05 neurônios estáticos e função de saída tipo tangente hiperbólica e a camada de saída com 01 neurônio estático e função de saída tipo linear. A arquitetura da rede com dinâmica interna consistiu de duas camadas intermediárias sendo a 1ª com 05 neurônios estáticos e a 2ª com 02 neurônios dinâmicos ambas com função de saída tipo tangente hiperbólica e

a camada de saída com 01 neurônio estático e função de saída tipo linear.

O treinamento da rede com dinâmica externa foi realizado em 52 iterações tendo alcançado um erro de 0,97, enquanto o treinamento da rede com dinâmica interna foi realizado em 1856 iterações tendo alcançado um erro de 0,99. As figuras 9 e 10 mostram o comportamento das RNA's após o treinamento.

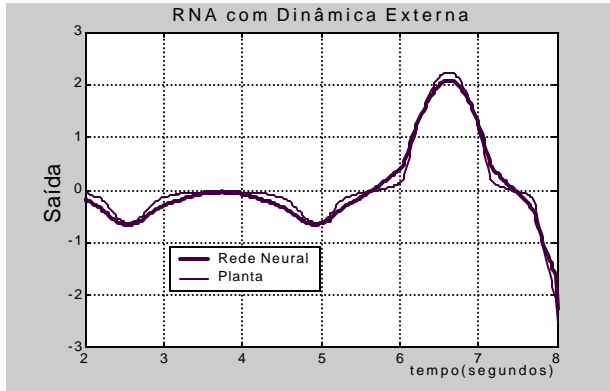


Figura 9: Resposta da RNA e da Planta ao sinal de treinamento

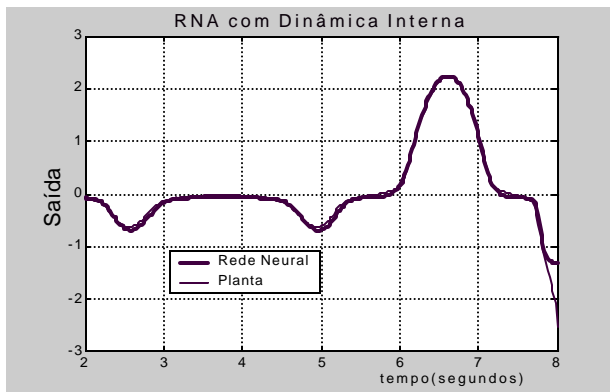


Figura 10: Resposta da RNA e da Planta ao sinal de treinamento

Para determinar a qualidade do modelo, utiliza-se testes de validação. Estes testes de validação do modelo são procedimentos criados para avaliar o modelo, independentemente das discrepâncias encontradas nos modelos neurais. Aqui será usado um conjunto de condições desenvolvido por Billings que faz medidas de funções de auto correlação do erro e correlações cruzadas como teste de validação do modelo neural [2][7]. O modelo é adequado se: (1) As estimativas de auto correlação próximas da origem apresentarem o formato de um impulso e todas as outras estiverem entre as linhas tracejadas e (2) Todas as estimativas de correlação cruzada estiverem entre as linhas tracejadas. A região definida entre as linhas tracejadas é a banda de confiança.

As figuras 10 a 20 representam resultados das funções correlações, calculadas no final do treinamento das duas redes utilizadas neste trabalho.

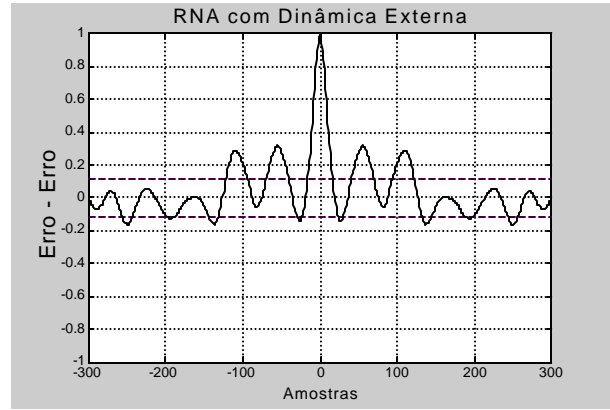


Figura 11: Função Correlação $\Phi_{ee}(t)$

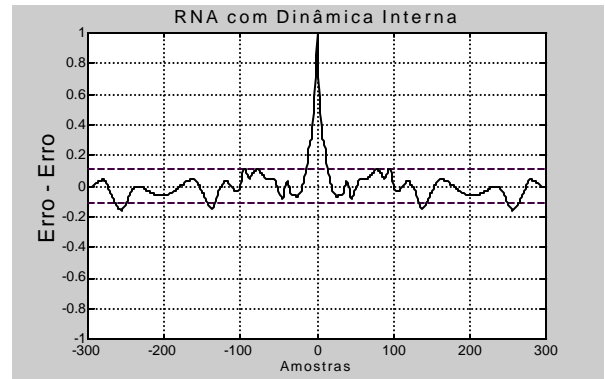


Figura 12: Função Correlação $\Phi_{ee}(t)$

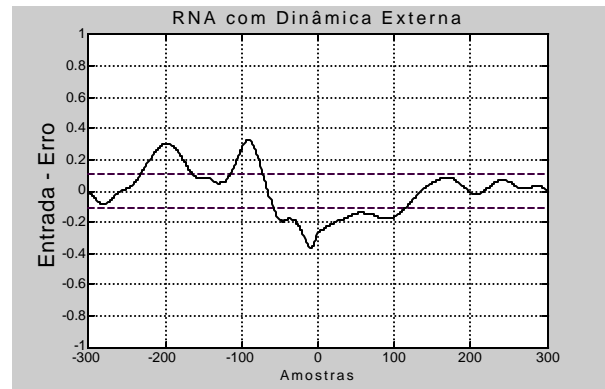


Figura 13: Função Correlação $\Phi_{ue}(t)$

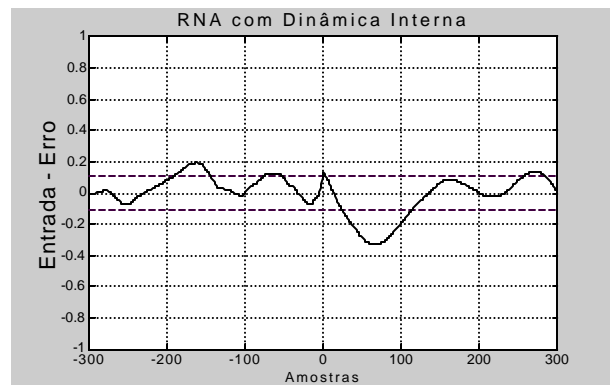


Figura 14: Função Correlação $\Phi_{ue}(t)$

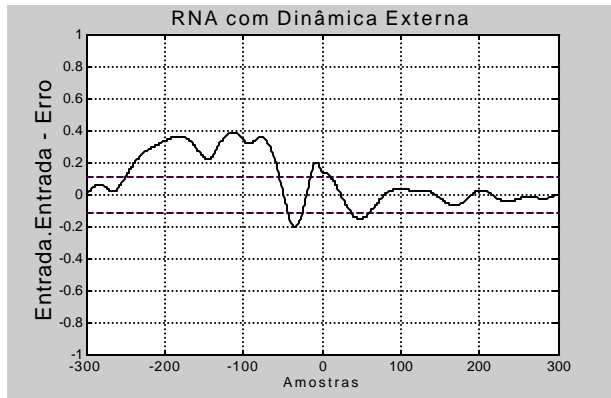


Figura 15: Função Correlação $\Phi_{u^2 e}(t)$

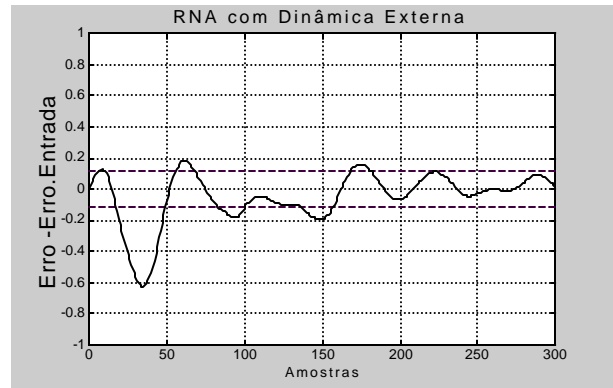


Figura 19: Função Correlação $\Phi_{e(a)}(t)$

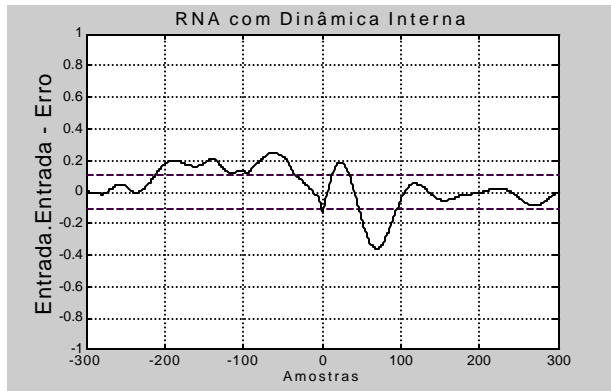


Figura 16: Função Correlação $\Phi_{u^2 e}(t)$

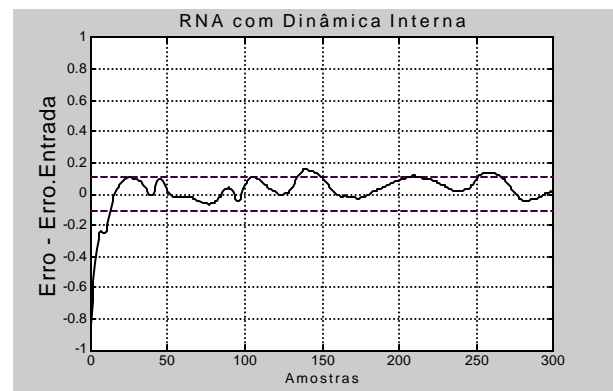


Figura 20: Função Correlação $\Phi_{e(a)}(t)$

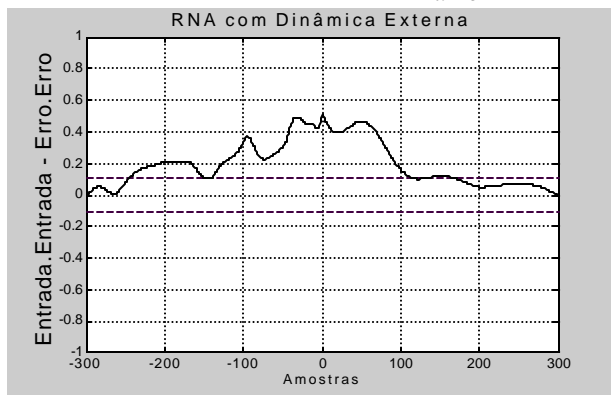


Figura 17: Função Correlação $\Phi_{u^2 e^2}(t)$

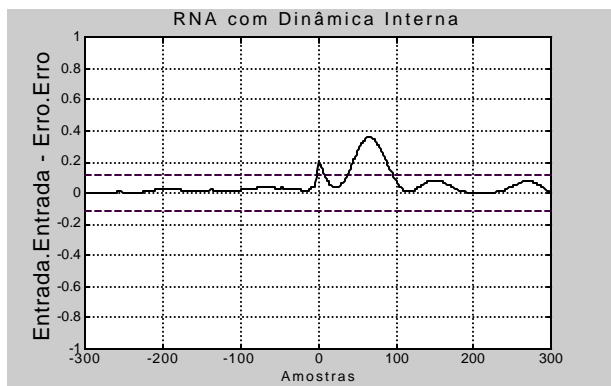


Figura 18: Função Correlação $\Phi_{u^2 e^2}(t)$

7. Conclusões

A utilização de Redes Neurais Artificiais na solução de problemas de Identificação de processos não lineares tem sido uma alternativa bastante promissora e explorada na literatura. Contudo, estabelecer a melhor arquitetura da Rede neural que será adotada para uma aplicação genérica, não constitui uma tarefa trivial. Neste trabalho, buscou-se avaliar de forma comparativa duas diferentes arquiteturas para o problema específico de identificação de um processo real não linear da classe de sistemas orientados a blocos. Para tal, utilizou-se uma RNA com dinâmica externa e portanto possuindo apenas neurônios estáticos, e uma RNA com dinâmica interna, apresentando 02 neurônios dinâmicos, baseada no modelo Hammerstein.

Embora o treinamento da RNA com dinâmica externa, tenha sido mais simples e com menor número de iterações, a análise dos resultados experimentais apresentados, mostra que a RNA com dinâmica interna apresenta melhor desempenho no que diz respeito a melhor representar a dinâmica da planta. Esse resultado reforça a idéia de que a melhor arquitetura de uma Rede Neural Identificadora é aquela que se aproxima mais da estrutura da planta. Dessa forma, toda e qualquer informação que se tenha a respeito da planta que se deseja identificar, deve ser incorporada no projeto da RNA Identificadora, garantindo melhores resultados.

Referências

- [1] Sjöberg J. (1995). Non-Linear System Identification with Neural Networks. Ph.D Thesis, Department of Electrical Engineering, Linköping University, Sweden.
- [2] Billings, S. A. and S. Chen (1995). Chaper 11- Neural Networks and System Identification. In G. W. Irwin, K. Warwick and K. J. Hunt (ed.), *Neural Networks Application in Control*. pp. 229—251, IEE Control Engineering Series 53, The Institution of Electrical Engineers, London, United Kingdom.
- [3] Al-Duwaish, H., M. N. Karim and V. Chandrasekar (1996). Use of Multilayer Feedforward Neural Networks in Identification and Control of Wiener Model. *IEE Proc.-Control Theory Appl.*, Vol. 143, No. 3, pp. 255—258.
- [4] U. Levin and K. S. Narendra. Control of Nonlinear Dynamical Systems using Neural Networks - Observability, Identification and Control. *IEEE Trans. on Neural Networks*, 7(1):30-42, 1996.
- [5] R. C. L. de Oliveira, (1999). Rede Neural com Dinâmica Interna Aplicada a Problema de Identificação e Controle Não Linear. Tese de Doutorado - Programa de Pós Graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.
- [6] F. M. De Azevedo. Uma Proposta de Modelos Formais de Neurônios e Redes Neurais. *Anais do III Congresso Brasileiro de Redes Neurais*, Florianópolis, SC, páginas 140-145, 1997.
- [7] R. C. L. de Oliveira, F. M. de Azevedo, e J. M. Barreto. Rede Neural com Dinâmica Linear Representada no Espaço de Estados. *Anais do III Congresso Brasileiro de Redes Neurais*, Florianópolis, SC, páginas 115-120, 1997.
- [8] Neural Networks Toolbox - The Math Works Inc., 1992.