

ABORDAGEM DE REDE NEURAL DE WAVELETS PARA IDENTIFICAÇÃO NÃO-LINEAR: PROJETO E ESTUDO DE CASOS DE LORENZ E BOX-JENKINS

Leandro dos Santos Coelho

Pontifícia Universidade Católica do Paraná, Programa de Pós-Graduação em Engenharia de Produção e Sistemas (PPGEPS)
Rua Imaculada Conceição, 1155 — Bairro Prado Velho — Curitiba — PR
leandro.coelho@pucpr.br

Viviana Cocco Mariani

Pontifícia Universidade Católica do Paraná, Programa de Pós-Graduação em Engenharia Mecânica (PPGEM)
Rua Imaculada Conceição, 1155 — Bairro Prado Velho — Curitiba — PR
viviana.mariani@pucpr.br

Resumo. Uma rede de wavelet é uma estrutura de regressão não-linear que representa mapeamentos de entrada e saída como superposição de versões dilatadas e transladadas de um função simples, que é localizada nos domínios do espaço e frequência. A rede de wavelet tem sido introduzida como uma rede neural feedforward especial, suportada pela teoria dos wavelets. Neste artigo, uma rede neural baseada em wavelet é introduzida para a identificação não-linear de sistemas dinâmicos. A rede neural baseada em wavelet com estrutura de modelo série-paralelo é avaliada para dois problemas: (i) previsão do sistema caótico de Lorenz, e (ii) previsão da concentração de CO₂ dos dados do forno de gás de Box e Jenkins. Os resultados de simulação e análise estatística demonstram a eficiência na previsão de um passo a frente (previsão de curtíssimo prazo) e as propriedades de aproximação da metodologia proposta de previsão de séries temporais.

Palavras chave: sistemas térmicos, identificação de sistemas térmicos, redes neurais de wavelets.

1. Introdução

A análise de sistemas dinâmicos é uma área de pesquisa emergente nos últimos anos e abrange as teorias de bifurcações, atratores, acoplamentos entre variáveis, complexidade, autômatos, previsão de séries temporais, sistemas não-lineares, inteligência artificial, otimização de sistemas, controle de sistemas, entre outras (Chen, 1999; Wang & Wang, 1999; Brown *et al.*, 2001). Em muitos casos, a aparente aleatoriedade em séries temporais pode ser devido ao comportamento caótico de um sistema não-linear (mas determinístico) e a sensibilidade do sistema caótico às condições iniciais impede a realização de previsões de séries temporais para um horizonte estendido.

Dada uma série de dados $\{x(t)\}_{t=0}^N \subset \mathfrak{R}$, a previsão de séries temporais equivale a inferir os valores futuros $\{x(t)\}_{t>N}$. Um algoritmo de previsão apropriado deve incorporar o comportamento dinâmico $x(t)$, ou seja, o mecanismo tão bom quanto possível de como os valores passados afetam os valores futuros. Quando o comportamento dinâmico pode ser assumido como linear, algoritmos bem consolidados na literatura de mínimos quadrados podem ser utilizados na estimativa de modelos matemáticos do tipo AR (*Auto Regressive*) ou ARMA (*Auto Regressive Moving Average*) (Johansson, 1993).

Quando não-linearidades estão presentes em um sinal, as capacidades de previsão dos algoritmos lineares são naturalmente limitadas. Uma classe desafiadora de problemas é a de previsão de séries temporais caóticas, isto é, quando o caos está presente no comportamento dinâmico do sinal. Um algoritmo eficiente para tal classe de problemas deve incorporar: (i) não-linearidade: porque sistemas não-lineares podem gerar caos, e (ii) ruído: porque alguns dados podem conter ruídos.

Entre as diversas metodologias possíveis e promissoras para capturar não-linearidades e incorporação de ruídos em problemas de identificação de sistemas não-lineares e/ou previsão de séries temporais estão as redes neurais (RNs). Uma definição de RN como “um sistema interconectado de funções parametrizadas” abrange muitos tipos de RNs e modelos de neurônios. A classificação das RNs é importante no contexto de comparar estruturas e melhor diferenciá-las. Um esquema de classificação simples pode ser realizado com algumas informações que descrevem uma RN, tais como (Redgers & Aleksander, 1995):

- (i) topologia: como as funções são interconectadas;
- (ii) arquitetura: o tipo e a utilização da RN;
- (iii) modelo do neurônio: o que as funções são;
- (iv) algoritmo de treinamento: como os parâmetros são configurados; e
- (v) escalonamento da operação: a sincronização das interações entre as funções.

A aproximação de funções, modelagem de sistemas, aprendizado de máquina e *data mining* são temas promissores para aplicações de RNs. A seleção das funções de ativação da rede neural é um fator crucial para a capacidade de aproximação e taxa de convergência de uma rede. Uma rede neural utilizada freqüentemente em previsão de séries temporais é o perceptron multicamadas (Rumelhart & McClelland, 1986), que na sua forma básica utiliza funções de ativação sigmoidais e treina os seus pesos com o algoritmo da retropropagação do erro (*backpropagation algorithm*). Devido a sua estrutura com múltiplas camadas, generalização global e a característica de subida de encosta do algoritmo de retropropagação do erro, o procedimento de treinamento desta rede neural freqüentemente se depara com armadilhas de mínimos locais da superfície de erro e apresenta convergência lenta.

Uma alternativa ao *perceptron* multicamadas, são as redes neurais com generalização local. Este caso ocorre em RNs quando poucos pesos afetam a resposta da RN, para os “pontos locais” no espaço de entrada. As RNs função de base radial (Orr, 1995), de *B-splines* (Harris *et al.*, 1996) e de *wavelets* (Zhang & Benveniste, 1992) são exemplos de RNs, com generalização local, em que a interferência no aprendizado é minimizada e o aprendizado é relativamente rápido, devido ao número mínimo de pesos ajustados a cada par de vetores de treinamento apresentado à RN.

Neste artigo aborda-se a fundamentação matemática e a utilização de rede neural de *wavelets* (RNWs) ou *wavenets* em dois estudos de caso de identificação (previsão de curtíssimo prazo com um passo à frente) de sistemas não-lineares. As RNWs são uma classe de redes neurais com uma camada intermediária, em que as funções de base são constituídas de uma família de funções de *wavelets*. As RNWs possuem propriedades similares às redes neurais de base radial e providenciam uma estrutura matemática eficiente para mapeamentos não-lineares. Uma motivação ao uso de RNWs deve-se aos resultados promissores de artigos apresentados na literatura em identificação de sistemas dinâmicos não-lineares e previsão de séries temporais usando RNWs (Szu *et al.*, 1993; Bakshi & Stephanopoulos, 1993; Chen & Lee, 1996; Hull & Pendse, 1997; Xu & Ho, 1999; Tan *et al.*, 2000; Rying *et al.*, 2002; Turkoglu *et al.*, 2003).

O artigo é organizado da seguinte forma. Na seção 2 são apresentadas noções, características e algoritmos para a concepção de RNWs. A aplicação de RNWs em identificação não-linear é apresentada na seção 3. Algumas simulações numéricas e o resultado da aplicação de RNWs são apresentados na seção 4. Finalizando, a conclusão e as perspectivas são apresentadas na seção 5.

2. Rede neural de *wavelets*

Nos últimos anos, os *wavelets* são um assunto muito atraente em muitas áreas de pesquisa científica e em engenharia. Especialmente, as RNWs inspiradas em redes neurais *feedforward* e decomposição de *wavelets* têm recebido considerável atenção (Zhang, 1997; Alonge *et al.*, 1999; Li *et al.*, 1999; Wei & Billings, 2002). Em contraste com as RNs baseadas nas clássicas funções de ativações sigmoidais, as RNWs possuem técnicas de concepção da estrutura da rede, aprendizado rápido e capacidades de análise com múltiplas resoluções (Rying *et al.*, 2002).

Recentemente, os pesquisadores têm começado a reconhecer que as RNWs tem algumas similaridades com sistemas nebulosos e rede neurais de função de base radial (Reyneri, 1999). Do ponto de vista de representação de funções, uma rede neural função de base radial é um esquema que representa uma função de interesse pela utilização de uma família de funções de base suportadas compactamente (ou localmente). A localidade das funções de base fazem as redes neurais de base radial serem mais apropriadas para o aprendizado de funções com variações locais e descontinuidades. Entretanto, as funções de base são pertencem a uma família de funções não ortogonais, e portanto, redundantes (Jiao *et al.*, 2001). Isto significa que para uma dada função, a representação através de uma rede neural função de base radial não é única e provavelmente não é a mais eficiente. Uma abordagem alternativa é substituir as funções de base da rede neural função de base radial por funções de base ortonormal, que consistem de funções de escala na teoria de *wavelets*. Estas redes são denominadas de redes neurais de *wavelets*. A utilização de redes neurais de *wavelets* ortonormais também facilita a análise teórica das propriedades de aproximação universal e consistência. Deve-se enfatizar que a estrutura de uma RNWs é similar a uma rede neural de base radial. No entanto, uma RNWs geralmente não é uma rede neural de base radial desde que as funções de escala multi-dimensionais podem não ser radialmente simétricas (Zhang *et al.*, 1995).

Em síntese, os *wavelets* oferecem uma estrutura matemática bem fundamentada para a representação de sinais (Daubechies, 1990; Nikolaou & Vuthandam, 1998). As RNWs, por outro lado, constituem-se em uma classe de aproximadores universais de funções (Kreinovich *et al.*, 1994) para identificação estrutural (modelos matemáticos caixa-cinza) de sistemas dinâmicos.

As RNWs foram originalmente propostas por Zhang & Benveniste (1992) e de forma similar a outras abordagens de redes neurais são utilizadas para representação precisa de mapeamentos de entrada e saída de sistemas dinâmicos, através do aprendizado de um conjunto de dados de treinamento. A representação de um mapeamento não-linear pelas RNWs depende da: (i) seleção apropriada da estrutura das RNWs; (ii) escolha da função de ativação na camada intermediária, ou seja, a função de base *wavenet*; (iii) escolha dos parâmetros de translação e dilatação (escala); e (iv) capacidade do algoritmo de aprendizado em determinar os parâmetros apropriados para a estrutura das RNWs selecionadas.

Uma RNWs pode ajustar os parâmetros da representação *wavelet* de um sistema dinâmico. No caso de um problema com N_i entradas é considerada a utilização de *wavelets* multidimensionais. Uma escolha simples é a realização do produto de N_i *wavelets* unidimensionais para cada entrada, empregando-se as equações

$$\Phi_j(x) = \prod_{k=1}^{N_i} \phi(z_{jk}) \quad (1)$$

com

$$z_{jk} = \frac{x(k) - m_{jk}}{d_{jk}}, \quad (2)$$

onde m_j e d_j são vetores de translação e dilatação, respectivamente. Neste estudo, consideram-se as RNWs com representação dada pela equação,

$$\hat{y} = \psi(x) = \sum_{j=1}^{N_w} c_j \Phi_j(x) + a_0 + \sum_{k=1}^{N_i} a_k x(k) . \quad (3)$$

A equação (3) pode ser vista como uma rede com N_i entradas, uma camada oculta de N_w *wavelets* de dimensão N_i com um termo de polarização (*bias*), a_0 , e um neurônio linear de saída. A RNWs pode ser representada pelo diagrama da Fig. 1. Neste artigo foi escolhida a primeira derivada de uma função gaussiana, $\phi(x) = \pm x \exp(\pm 0,5x^2)$, como a *wavelet*-mãe.

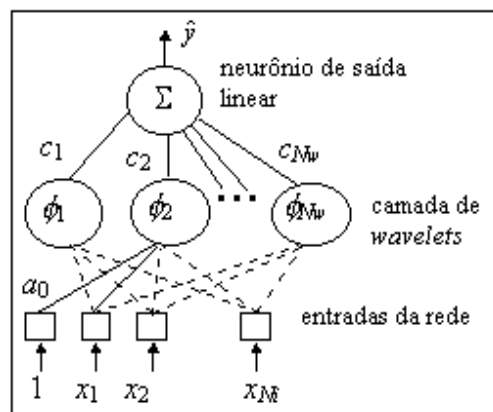


Figura 1. Estrutura de uma RNWs.

Na próxima seção são apresentados dois casos de sistemas não-lineares, onde serão utilizadas as RNWs na identificação destes sistemas dinâmicos.

3. Estudo de casos de identificação não-linear de sistemas

3.1. Estudo de caso 1: Atrator de Lorenz

Em 1963, o matemático e meteorologista Edward N. Lorenz (1963) descreveu uma família de equações diferenciais tridimensionais que representam um modelo simplificado para turbulência atmosférica debaixo do centro de um trovão. A solução destas equações, usando métodos numéricos, pode ser extremamente complicada. Estas equações não-lineares, conhecidas como equações ou atrator de Lorenz, são muito interessantes e estudadas como um exemplo de dinâmica caótica (Alligood *et al.*, 1996). As equações de Lorenz são:

$$\frac{dx}{dt} = \sigma(y - x) \quad (4)$$

$$\frac{dy}{dt} = rx - y - xz \quad (5)$$

$$\frac{dz}{dt} = xy - bz \quad (6)$$

onde σ, r e $b \in \mathfrak{R}_+$. Neste artigo é apresentado um estudo para $(\sigma, r, b) = (10; 28; 8/3)$. Neste caso, substituindo-se os valores de (σ, r, b) nas equações (4), (5) e (6) obtém-se

$$\frac{dx}{dt} = 10(y - x) \quad (7)$$

$$\frac{dy}{dt} = 28x - y - xz \quad (8)$$

$$\frac{dz}{dt} = xy - \frac{8}{3}z \quad (9)$$

O atrator caótico para este caso é apresentado na Fig. 2. Os correspondentes comportamentos dinâmicos nos planos x - y , y - z e z - x , são também apresentados, estes na Fig. 3.

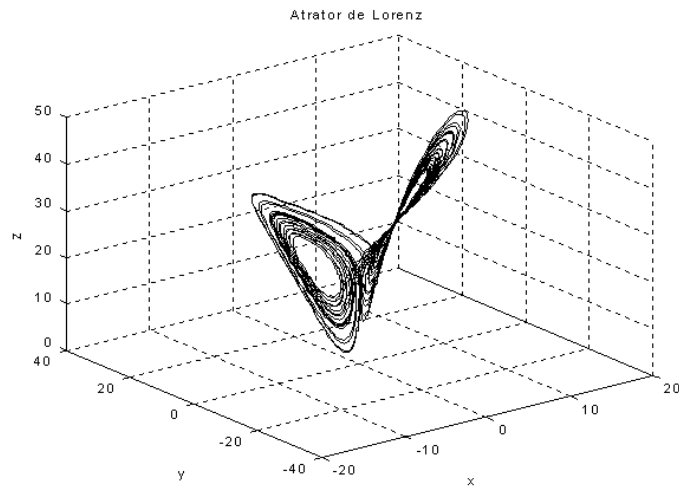
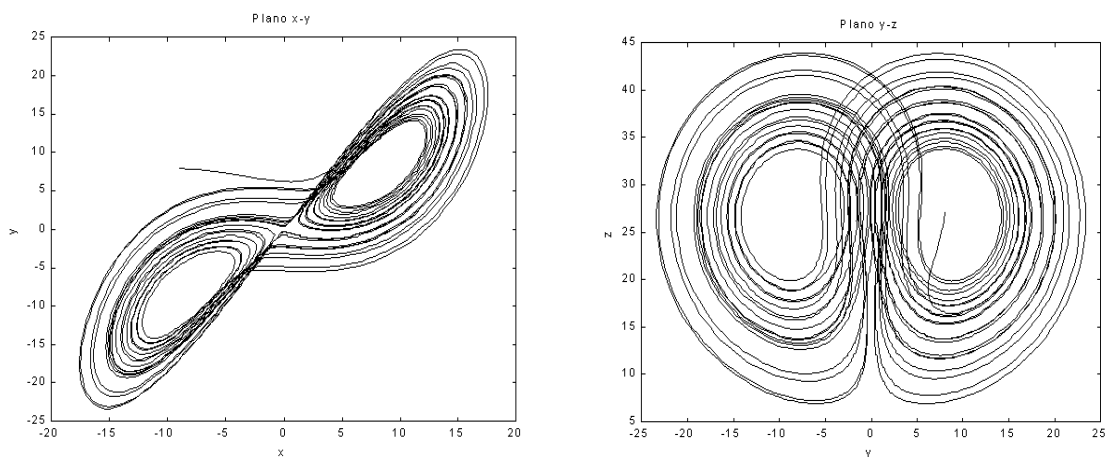


Figura 2. Atrator de Lorenz para $(\sigma, r, b) = (10; 28; 8/3)$.



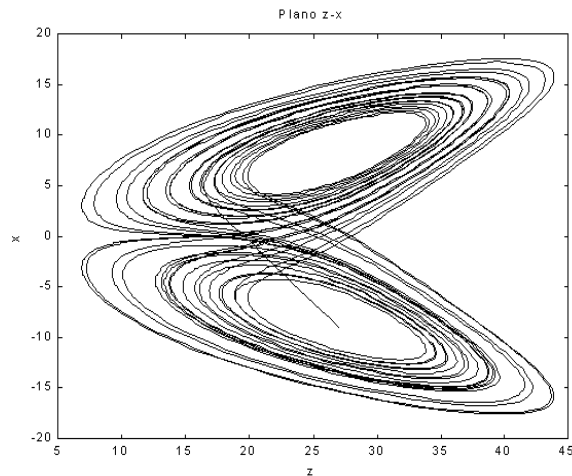


Figura 3. Comportamento dinâmico do atrator de Lorenz para $(\sigma, r, b) = (10; 28; 8/3)$.

Um conjunto de 5900 amostras é utilizado para identificação deste sistema, estes apresentados na Fig. 4, onde o objetivo é identificar o comportamento dinâmico da coordenada z do atrator de Lorenz.

Desde que as equações (7) a (9) representam um sistema dinâmico contínuo em vez de um sistema dinâmico discreto, é necessário utilizar um método de integração numérica. Neste caso, escolheu-se o método de Runge-Kutta de 4ª ordem com passo de 0,005 com condições iniciais $(x(0), y(0), z(0)) = (-9; 8; 27)$. O experimento foi configurado para utilizar 2900 amostras (amostras 1 a 2901) para a fase de estimação e 2900 amostras (amostras 2902 a 5902) para a fase de validação do modelo matemático, obtido através de uma RNWs.

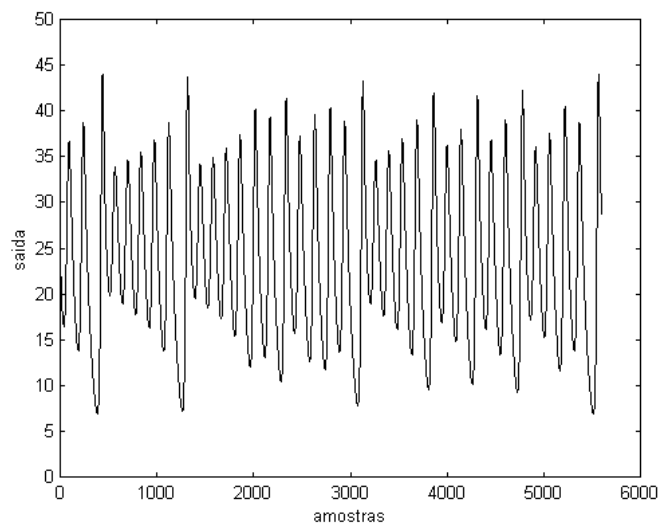


Figura 4. Série temporal da coordenada $z(t)$ usada para identificação do atrator de Lorenz para o caso de $(\sigma, r, b) = (10; 28; 8/3)$.

3.2 Estudo de caso 2: Forno de gás de Box e Jenkins

A fornalha de gás de Box e Jenkins (Box *et al.*, 1994) consiste de uma fornalha onde a entrada $u(t)$ é a taxa de alimentação de gás metano (pés cúbicos por minuto), e a saída $z(t)$ é a concentração de dióxido de carbono (%CO₂) em uma mistura de gases.

Um conjunto de 296 amostras (pares de dados de entrada e saída) é disponível para a identificação deste sistema. O experimento foi configurado para utilizar 200 amostras para a fase de estimação e 96 amostras para a fase de validação do modelo matemático, obtido através da RNWs. Os dados da entrada e da saída da fornalha de gás são apresentados na Fig. 5.

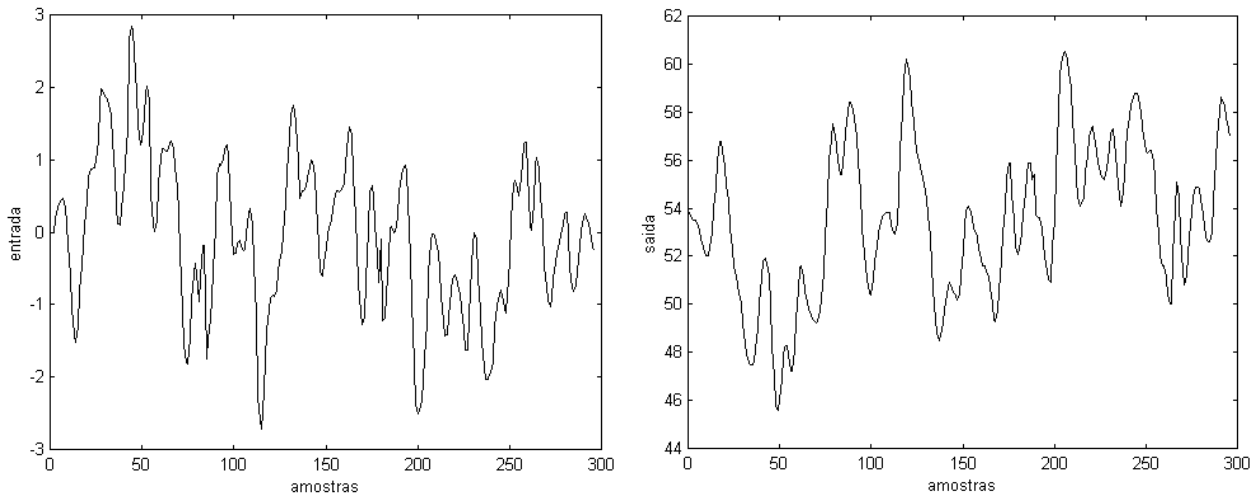


Figura 5. Dados de entrada e saída da fornalha de gás.

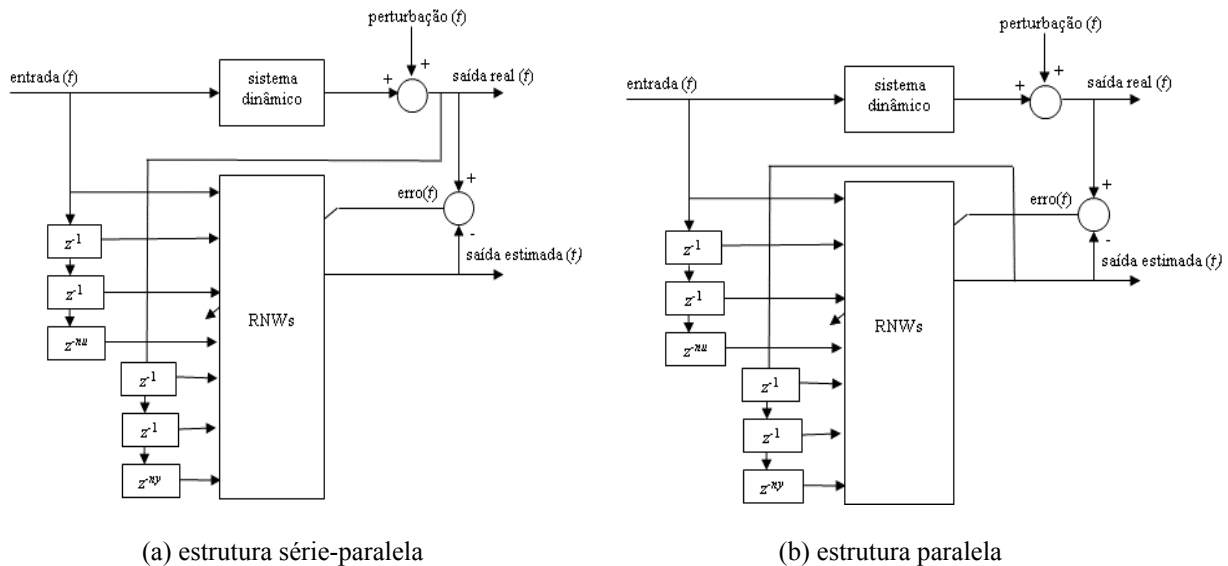
4. Resultados de simulação para o atrator de Lorenz e a fornalha de gás de Box e Jenkins

O procedimento de análise dos resultados obtidos nos problemas de identificação não-linear, é dividido nas seguintes etapas:

- (i) obtenção dos dados de entrada e saída do sistema dinâmico;
- (ii) escolha das entradas (variáveis a serem consideradas e seus respectivos atrasos de tempo) e estrutura da RNWs (número de *wavelets*, escalas e espaço de busca);
- (iii) determinação de um procedimento de otimização da RNWs adotando o algoritmo do gradiente estocástico (Zhang & Benveniste, 1992). A função objetivo a ser minimizada é regida pelo erro médio quadrático (MSE). Foram utilizadas 200 épocas para o treinamento das RNWs para os dois casos estudados;
- (iv) estimação dos parâmetros do modelo matemático (fase de estimação ou treinamento); e
- (v) validação do modelo matemático (fase de validação ou testes).

Para ajustar a RNWs para identificar o comportamento dinâmico dos sistemas não-lineares é necessário escolher a estrutura do modelo neural. Na estrutura série-paralela, conforme mostrado na Fig. 6(a), as saídas do sistema dinâmico são utilizadas como entradas para o modelo da RNWs. Neste caso somente um passo de previsão a frente é possível e a estrutura tem dinâmica externa. Em contrapartida, na estrutura paralela, conforme ilustrada na Fig. 6(b), a própria saída estimada da RNW é utilizada para criar entradas atrasadas no tempo para a RNWs. Esta estrutura pode ser considerada recorrente e com dinâmica interna. Além disso, esta estrutura pode obter previsões sobre um pequeno período de tempo.

Após a realização de estudos de viabilidade do projeto optou-se por uma estrutura série-paralela, conforme apresentado na Fig. 6(a), em vez da estrutura paralela, principalmente pela facilidade de sua implementação quando o objetivo é a previsão da saída do sistema real com um passo a frente, isto é, para o instante $t+1$. O modelo de previsão escolhido para este artigo é um modelo NAR (*Nonlinear AutoRegressive*) para o caso do atrator de Lorenz sem entradas $u(t)$, ou seja, $nu = 0$ e um modelo NARX (*Nonlinear AutoRegressive with exogenous inputs*) para o caso da fornalha de gás de Box e Jenkins.



(a) estrutura série-paralela (b) estrutura paralela
 Figura 6. Estruturas avaliadas da RNWs para previsão/identificação de sistemas dinâmicos.

O critério de desempenho avaliado para o sistema dinâmico a ser identificado é o coeficiente de correlação múltipla, R^2 , entre a saída real, $z(t)$, e a saída estimada, $\hat{z}(t)$, dado por

$$R^2 = 1 - \frac{\sum_{t=1}^N [z(t) - \hat{z}(t)]^2}{\sum_{t=1}^N (z(t) - \bar{z})^2}, \tag{10}$$

onde $z(t)$ é a saída do processo, \bar{z} é a média das N amostras medidas para a saída do processo. Quando o valor de R^2 é igual a 1,0 indica uma adequação exata do modelo para os dados medidos do processo. O valor de R^2 entre 0,9 e 1,0 é considerado suficiente para aplicações práticas, principalmente em identificação não-linear e em sistemas de controle baseados em modelo (Schaible *et al.*, 1997). Os resultados para as RNWs são baseados no coeficiente de correlação múltipla, R^2 , e estes são apresentados nas Tabs. 1 e 2, respectivamente.

Tabela 1. Resultados obtidos através da RNWs para o estudo de caso 1 (atrator de Lorenz) com 200 épocas usadas na fase de treinamento.

entradas da RNWs	wavelets na camada oculta	R^2 (estimação)	R^2 (validação)	R^2 (total)
[$z(t-1)$ $z(t-2)$]	2	0,987648	0,987756	0,987705
[$z(t-1)$ $z(t-2)$]	3	0,991579	0,999999	0,991580
[$z(t-1)$ $z(t-2)$]	4	0,904125	0,913163	0,908787
[$z(t-1)$ $z(t-2)$]	5	0,908747	0,917281	0,913149
[$z(t-1)$ $z(t-2)$ $z(t-3)$]	2	0,981019	0,980813	0,980916
[$z(t-1)$ $z(t-2)$ $z(t-3)$]	3	0,980784	0,980640	0,980713
[$z(t-1)$ $z(t-2)$ $z(t-3)$]	4	0,981023	0,980763	0,980892
[$z(t-1)$ $z(t-2)$ $z(t-3)$]	5	0,980827	0,980635	0,980732

Tabela 2. Resultados obtidos através da RNWs para o estudo de caso 2 (fornalha de gás) com 200 épocas usadas na fase de treinamento.

entradas da RNWs	wavelets na camada oculta	R^2 (estimação)	R^2 (validação)	R^2 (total)
[$u(t)$ $y(t)$]	2	0,987441	0,942387	0,981012
[$u(t)$ $y(t)$]	3	0,988065	0,942489	0,981426
[$u(t)$ $y(t)$]	4	0,987436	0,964127	0,985176
[$u(t)$ $y(t)$]	5	0,986674	0,970486	0,985876
[$u(t)$ $y(t)$ $y(t-1)$]	2	0,987755	0,940664	0,981095
[$u(t)$ $y(t)$ $y(t-1)$]	3	0,987322	0,961891	0,984739
[$u(t)$ $y(t)$ $y(t-1)$]	4	0,987493	0,961681	0,984811
[$u(t)$ $y(t)$ $y(t-1)$]	5	0,988007	0,960951	0,985009

Os resultados de identificação através de RNWs foram precisos para a identificação dos casos apresentados. Na Fig. 7 são apresentados os gráficos relacionados a saída estimada e ao erro de estimação do melhor resultado, apresentado na Tab. 1, obtido com a RNWs para a identificação do atrator de Lorenz. O melhor resultado foi com a utilização de 3 *wavelets* na camada oculta da RNWs. Neste caso, o coeficiente de correlação múltipla, R^2 , ficou próximo da unidade nas fases de estimação e validação, mostrando a precisão e robustez da RNWs na identificação um passo a frente, $\hat{z}(t+1)$, do atrator de Lorenz. Além disso, o erro máximo, o médio e o desvio padrão do erro da RNWs foram de [0,057867 -0,000367 0,012621], respectivamente. Neste caso, o erro absoluto normalizado foi de 0,001484% e observa-se que o erro médio tem característica de ruído Gaussiano decorrelacionado.

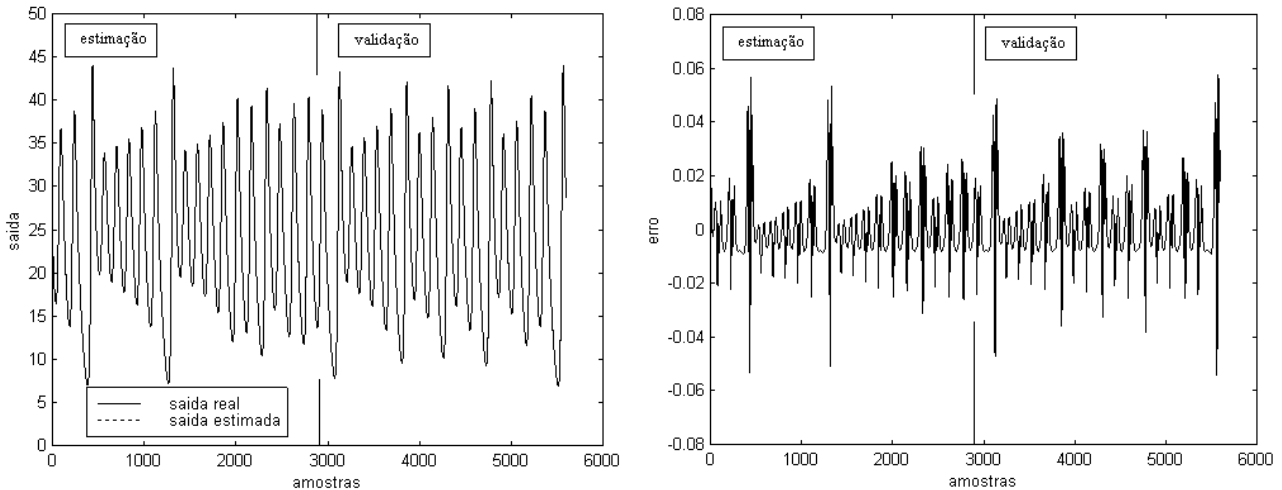


Figura 7. Resultados da previsão do atrator de Lorenz com 3 *wavelets* na camada oculta da RNWs.

Em relação ao caso da fornalha de gás, o melhor resultado foi com uma RNWs com 5 *wavelets* na camada oculta, este apresentado na Fig. 8. Nota-se que no caso da fornalha de gás os resultados com 2 a 5 *wavelets* na camada oculta da RNWs ficaram próximos. O melhor resultado da RNWs foi para erro máximo, o erro médio e o desvio padrão do erro com valores [1,760198 0,088400 0,41665619], respectivamente. Notou-se que o erro máximo ocorreu na fase de validação, mostrando a necessidade de estudos posteriores para melhorar as propriedades de generalização da RNWs para o caso da fornalha de gás.

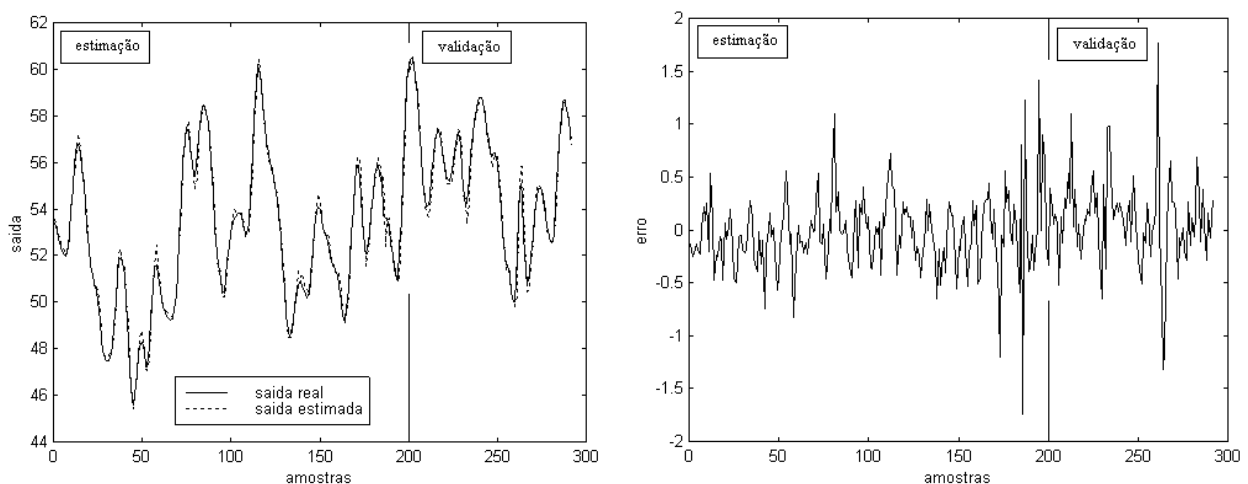


Figura 8. Resultados da previsão da fornalha de gás de Box e Jenkins com 5 *wavelets* na camada oculta da RNWs.

Em síntese, os resultados obtidos mostram a robustez da RNWs na previsão de curto prazo para sistemas dinâmicos não-lineares. Entretanto, alguns questionamentos podem ser realizados. A solução do problema de seleção estrutural que está presente em RNWs depende dos valores atribuídos a dilatações e translações das *wavelets*. Neste caso, existe

um compromisso entre o número de *wavelets* selecionados (complexidade computacional) e a qualidade de aproximação da RNWs para a previsão de séries temporais com comportamento dinâmico não-linear.

5. Conclusão

As RNWs possuem as vantagens da alta resolução do *wavelets* e as potencialidades de aprendizado das redes neurais *feedforward*. Além disso, as RNWs apresentam vantagens devido à capacidade de aproximação de funções não-lineares, rapidez e eficiência do aprendizado, possibilitando a obtenção de resultados precisos, oferecendo soluções eficientes para problemas de aproximação multidimensional. Estas características motivaram este artigo a apresentar os fundamentos e a aplicação de RNWs para o aprendizado de sistemas com dinâmica complexa.

Neste artigo foram analisados dois casos visando à previsão de curto prazo (um passo a frente) através de simulações computacionais para: (i) o atrator de Lorenz, e (ii) a concentração de CO₂ dos dados do forno de gás de Box e Jenkins. Os resultados de simulação mostraram que as RNWs apresentaram aprendizado e convergência rápidos e uma boa aproximação da dinâmica não-linear para os dois casos abordados. Estes resultados confirmam as propriedades de análise multiresolução das RNWs de aproximarem sistemas que possuem não-linearidade e/ou dinâmica complexa. No caso do atrator de Lorenz os resultados da RNWs ficaram entre $0,904125 \leq R^2 \leq 0,999999$, e no caso da fornalha da gás os resultados obtidos foram entre $0,942387 \leq R^2 \leq 0,988065$, o que mostram serem precisos e suficientes para aplicações futuras em algoritmos de controle baseados em modelo, tais como algoritmos de controle adaptativo indireto, controle preditivo e controle neural.

Alguns experimentos serão realizados em trabalhos futuros de forma a aprimorar os estudos e resultados de simulação apresentados neste artigo. Entre os experimentos os autores desejam realizar:

- (i) enfatizar características locais da identificação não-linear. As redes neurais têm capacidade limitada para enfatizar características locais, tais como descontinuidades na curvatura e variações bruscas em valores. Isto é devido a função objetivo que geralmente guia a construção das RNWs a serem otimizadas baseando-se em erro médio quadrático global, assim a qualidade da identificação das características locais não é enfatizada;
- (ii) um estudo comparativo de precisão e complexidade computacional de modelos NAR e NARX em relação aos modelos matemáticos do tipo NARMA (*Nonlinear AutoRegressive Moving Average*) e NARMAX (*Nonlinear AutoRegressive Moving Average with eXogenous inputs*);
- (iii) uma análise de auto-correlação e correlação cruzada entre variáveis de entrada(s), saída(s) e erro(s) da RNWs; e
- (iv) um estudo das dificuldades de identificação não-linear para conjuntos de dados não-estacionários, ruidosos e com comportamento caótico para sistemas dinâmicos MIMO (*Multi-Inputs Multi-Outputs*).

6. Referências

- Alligood, K. T., Sauer, T. D. and Yorke, J. A., 1996, "Chaos: an Introduction to Dynamical Systems", Springer, London.
- Alonge, F., D'Ippolito, F., Mantione, S. and Raimondi, F. M., 1999, "A New Method for Optimal Synthesis of Wavelet-based Neural Networks Suitable for Identification Purposes", 14th IFAC World Conference, Beijing, P. R. China, pp. 445-450.
- Bakshi, B. R. and Stephanopoulos, G., 1993, "Wave-net: a Multiresolution, Hierarchical Neural Network with Localized Learning", AIChE Journal, Vol. 39, No. 1, pp. 57-81.
- Box, G., Jenkins, G. and Reinsel, G., 1994, "Time Series Analysis, Forecasting and Control", 3rd edition, Prentice Hall, Englewood Cliffs, NJ.
- Brown, R., Berezdivin, R. and Chua, L. O., 2001, "Chaos and Complexity", International Journal of Bifurcation and Chaos, Vol. 11, No. 1, pp. 19-26.
- Chen, C. H. and Lee, G. G., 1996, "Multiresolution Wavelet Analysis Based Feature Extraction for Neural Network Classification", IEEE International Conference on Neural Networks, Washington, DC, Vol. 3, pp. 1416-1421.
- Chen, G. (ed.), 1999, "Controlling Chaos and Bifurcations in Engineering Systems", CRC Press, Boca Raton, FL.
- Daubechies, I., 1990, "The Wavelet Transform, Time-frequency Localization and Signal Analysis", IEEE Transactions on Information Theory, Vol. 36, No. 5, pp. 961-1005.
- Harris, C. J., Brown, M., Bossley, K. M., Mills, D. J. and Ming, F., 1996, "Advances in Neurofuzzy Algorithms for Real-time Modeling and Control", Engineering Applications of Artificial Intelligence, Vol. 9, No. 1, pp. 1-16.
- Jiao, L., Pan, J. and Fang, Y., 2001, "Multiwavelet Neural Network and Its Approximation Properties", IEEE Transactions on Neural Networks, Vol. 12, No. 5, pp. 1060-1066.
- Johansson, R., 1993, "System Modeling and Identification", Prentice-Hall, Englewood Cliffs, NJ.
- Li, X., Wang, Z., Xu, L., Liu, J., 1999, "Combined Construction of Wavelet Neural Networks for Nonlinear System Modeling", 14th IFAC World Conference, Beijing, P. R. China, pp. 451-456.
- Lorenz, E. N., 1963, "Determinist Nonperiodic Flow", J. Atmos. Science, Vol. 20, pp. 130.

- Kreinovich, V.; Sirisaengtaksin, O. and Cabrera, S., 1994, "Wavelet Neural Networks are Asymptotically Optimal Approximators for Functions of One Variable", Proceedings of IEEE International Conference on Neural Networks, Orlando, FL, pp. 299-304.
- Nikolaou, M. and Vuthandam, P., 1998, "FIR Model Identification: Parsimony through Kernel Compression with Wavelets", AICHE Journal, Vol. 44, No. 1, pp. 141-150.
- Orr, M., 1995, "Regularisation in the Selection of Radial Basis Function Centres", Neural Computation, Vol. 7, No. 3, pp. 606-623.
- Redgers, A. and Aleksander, I., 1995, "Digital Neural Networks", In: Irwin, G. W.; Warwick, K.; Hunt, K. J. (eds.). Neural Network Applications in Control, IEE Control Engineering Series 53, The Institution of Electrical Engineers. Chapter 2, pp. 17-32.
- Reyneri, L. M., 1999, "Unification of Neural and Wavelet Networks and Fuzzy Systems", IEEE Transactions on Neural Networks, Vol. 10, No. 4, pp. 801-814.
- Rumelhart, D. E. and McClelland, J. L., 1986, "Parallel Distributed Processing", Cambridge, MA: MIT Press.
- Rying, E. A., Bilbro, G. L. and Lu, J. -C., 2002, "Focused Local Learning with Wavelet Neural Networks", IEEE Transactions on Neural Networks, Vol. 13, No. 2, pp. 304-319.
- Schaible, B.; Xie, H. and Lee, Y. C., 1997, "Fuzzy Logic Models for Ranking Process Effects", IEEE Transactions on Fuzzy Systems, Vol. 5, No. 4, pp. 545-556.
- Szu, H. H., Yang, X. -Y. and Sheng, Y., 1993, "Neural Network and Wavelet Transform for Scale-invariant Data Classification", Physica Review E, Vol. 48, No. 2, pp. 1497-1501.
- Tan, Y., Dang, X., Liang, F. and Su, S. -Y., 2000, "Dynamic Wavelet Neural Network for Nonlinear Dynamic System Identification", Proceedings of the IEEE International Conference on Control Applications, Anchorage, AL, pp. 214-219.
- Turkoglu, I., Arslan, A. and Ilkay, E., 2003, "A Wavelet Neural Network for the Detection of Heart Valve Diseases", Expert Systems, Vol. 20, No. 1, pp. 1-7.
- Wang, J. and Wang, X., 1999, "A Global Control of Polynomial Chaotic Systems", International Journal of Control, Vol. 72, No. 10, pp. 911-918.
- Wei, H. -L. and Billings, S. A., 2002, "Identification of Time-varying Systems Using Multiresolution Wavelet Models", International Journal of Systems Science, Vol. 33, No. 15, pp. 1217-1228.
- Xu, J. and Ho, D. W. C., 1999, "Adaptive Wavelet Networks for Nonlinear System Identification", Proceedings of the American Control Conference, San Diego, CA, pp. 3472-3473.
- Zhang, J., Walter, G. G., Miao, Y. and Lee, W. N. W., 1995, "Wavelet Neural Networks for Function Learning", IEEE Transactions on Signal Processing, Vol. 43, No. 6, pp. 1485-1497.
- Zhang, Q., 1997, "Using Wavelet Network in Nonparametric Estimation", IEEE Transactions on Neural Networks, Vol. 8, No. 2, pp. 227-236.
- Zhang, Q. and Benveniste, A., 1992, "Wavelet Networks", IEEE Transactions on Neural Networks, Vol. 3, No. 6, pp. 889-898.

WAVELET-BASED NEURAL NETWORK APPROACH IN NONLINEAR IDENTIFICATION: DESIGN AND LORENZ AND BOX-JENKINS CASES STUDY

Leandro dos Santos Coelho
Pontifícia Universidade Católica do Paraná
Programa de Pós-Graduação em Engenharia de Produção e Sistemas (PPGEPS)
Rua Imaculada Conceição, 1155 — Bairro Prado Velho — Curitiba — PR, Brazil
leandro.coelho@pucpr.br

Viviana Cocco Mariani
Pontifícia Universidade Católica do Paraná
Programa de Pós-Graduação em Engenharia Mecânica (PPGEM)
Rua Imaculada Conceição, 1155 — Bairro Prado Velho — Curitiba — PR, Brazil
viviana.mariani@pucpr.br

Abstract

A wavelet network is a nonlinear regression structure that represents input-output mappings as the superposition of dilated and translated versions of a single function, which is localized both in the space and frequency domains. The wavelet network has been introduced as a special feedforward neural network supported by the wavelets theory. In this paper, a wavelet-based neural network is introduced for nonlinear identification of dynamic systems. The wavelet-based neural network with serie-parallel structure is evaluated on two cases: (i) prediction of Lorenz chaotic system, and (ii) prediction of CO₂ concentration of Box and Jenkins gas furnace data. Simulation results and statistical analysis demonstrate the effectiveness and the approximation properties of the proposed methodology for prediction time series with forecasting of one step ahead.

Keywords: thermal systems, identification of thermal systems, wavelet-based neural network.