

Aprendizado de Dinâmica Caótica usando Redes Neurais de *Wavelets*

Leandro dos Santos Coelho

Pontifícia Universidade Católica do Paraná

Grupo Produtrônica, Programa de Pós-Graduação em Engenharia de Produção e Sistemas

Laboratório de Automação e Sistemas, LAS/PPGEPS/CCET/PUCPR

Rua Imaculada Conceição, 1155, CEP 80215-901, Curitiba, PR, Brasil

E-mails: lscoelho@rla01.pucpr.br

Abstract

In this paper, a wavelet-based neural network is introduced for nonlinear identification of dynamic systems with chaotic behavior (chaotic time series). The structure of the wavelet-based neural network is similar to that of radial basis function neural network, except that there the activation function of the hidden nodes is replaced by wavelet functions. The proposed wavelet-based neural network is evaluated on two case studies: (i) Hénon map, and (ii) Rössler system. Simulation results demonstrate the effectiveness of the proposed identification methodology with wavelet-based neural network.

1. Introdução

Os sistemas dinâmicos caóticos são um tópico de diversas pesquisas nos últimos anos, pois abrange as teorias de bifurcações, atratores, acoplamentos entre variáveis, previsão de séries temporais, sistemas não-lineares, controle de sistemas, entre outras [1]-[3]. 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.

Entre as diversas metodologias possíveis para a identificação de sistemas não-lineares e/ou previsão de séries temporais estão as redes neurais. Um tipo de rede neural com propriedades interessantes para estas aplicações é a rede neural de *wavelets* (*RNWs*) ou *wavenets*. 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 as redes neurais de base radial e providenciam uma estrutura matemática eficiente para mapeamentos não-lineares.

Neste artigo aborda-se a fundamentação matemática e a utilização de *RNWs* em dois estudos de caso de identificação de sistemas com comportamento caótico. Uma motivação ao uso de *RNWs* é devido aos resultados promissores de artigos apresentados na literatura em identificação de sistemas dinâmicos não-lineares usando esta rede neural [4]-[14].

O artigo é organizado da seguinte forma. Na seção 2 são apresentados 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 (*RNWs*)

Os *wavelets* oferecem uma estrutura matemática bem fundamentada para a representação de sinais [10]-[11]. As *RNWs*, por outro lado, constituem-se em uma classe de aproximadores universais de funções [7] para identificação estrutural (modelos matemáticos caixa-cinza) de sistemas.

As *RNWs*, de forma similar a outras abordagens de neurais neurais, são utilizadas para representar de forma precisa 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 pela *RNWs* depende da: (i) seleção apropriada da estrutura da *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 da *RNWs* selecionada.

Uma *RNW* pode ajustar os parâmetros da representação *wavelet* de um sistema dinâmico. Entretanto, um procedimento de pré-processamento dos dados e o conhecimento *a priori* de parâmetros para o projeto da *RNW* são necessários. Na abordagem apresentada neste artigo a estrutura da *RNW* é ajustada através de um procedimento de ortogonalização de Gram-Schmidt [12].

A investigação da estrutura de uma *RNWs*, conforme mencionado na introdução deste artigo, revela semelhanças com a arquitetura de uma rede neural de base radial, esta uma rede neural com três camadas (entrada, intermediária e saída).

No caso de um problema com N_i entradas pode-se considerar a utilização de *wavelets* multidimensionais. Uma escolha simples é a realização do produto de N_i *wavelets* monodimensionais para cada entrada, empregando-se as seguintes equações:

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

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, considera-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 *RNW* pode ser representada pelo diagrama da figura 1. Neste artigo foi escolhida a primeira derivada de uma função gaussiana, $\phi(x) = \pm x \exp(\pm 0,5x^2)$, como uma *wavelet*-mãe.

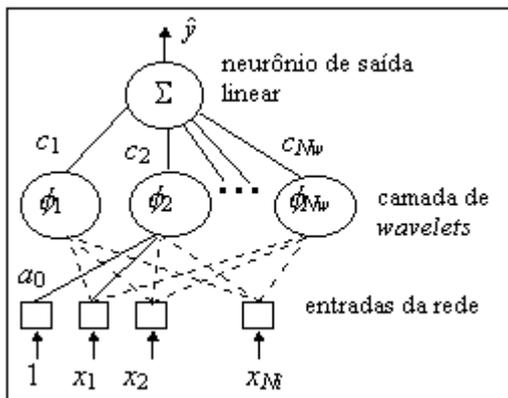


Figura 1. Estrutura de uma *RNWs*.

Na próxima seção são apresentados dois estudos de caso de sistemas caóticos visando a utilização de *RNWs* para a identificação não-linear destes sistemas dinâmicos.

3. Estudos de caso para identificação não-linear usando *RNWs*

3.1. Estudo de Caso 1: Mapa de Hénon

O astrônomo francês M. Hénon, em 1975, mostrou um interessante fenômeno (os mapas de Hénon), presente em mapas de Poincaré gerados por equações diferenciais. O mapa de Hénon [13] é dado pela seguinte equação a diferenças:

$$x(t+1) = 1 - a[x(t)]^2 + 0,3x(t-1). \quad (4)$$

Pode-se observar que dependendo do valor escolhido para o parâmetro a , o sistema dinâmico exibe diferentes comportamentos, desde regimes periódicos até comportamento de caos determinístico. Neste caso, o sistema dinâmico é sensível a pequenas perturbações tornando uma tarefa complexa à análise do modelo matemático que representa o mapa de Hénon.

Neste artigo considerou-se uma série temporal caótica composta de 400 valores gerados a partir das condições iniciais $x(0)=0,1$ e $x(1)=0,3$ para o parâmetro $a=1,4$. As simulações foram configuradas para utilizar-se 200 amostras para a fase de estimação e 200 amostras, para a fase de validação do modelo obtido pela *RNWs* para o mapa de Hénon.

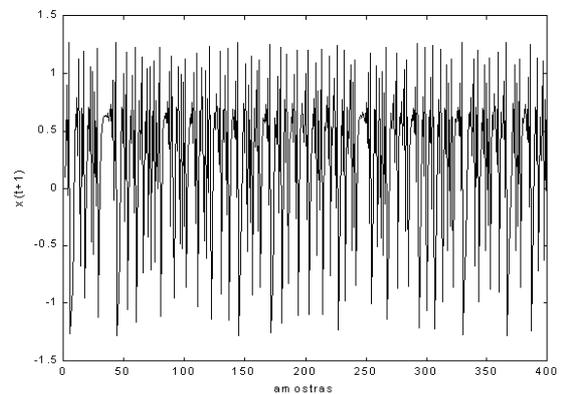


Figura 2. Dados usados na previsão do mapa de Hénon via *RNWs*.

3.2. Estudo de Caso 2: Sistema de Rössler

O cientista alemão O. Rössler [14] propôs um atrator caótico constituído de equações diferenciais não-lineares. As equações de Rössler são [15]:

$$\frac{dx}{dt} = -y - z, \quad (5)$$

$$\frac{dy}{dt} = x + ay, \quad (6)$$

$$\frac{dz}{dt} = b + (x - c)z, \quad (7)$$

que exibem comportamento caótico para $(a, b, c) = (0,36; 0,40; 4,50)$. O atrator caótico para este caso é apresentado na figura 3. Os correspondentes comportamentos dinâmicos nos planos x - y , y - z , e z - x , são também apresentados, estes na figura 4.

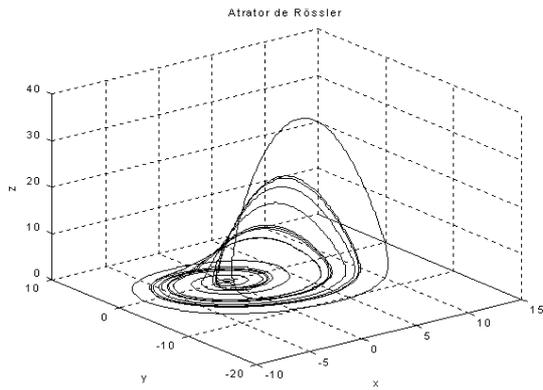


Figura 3. Atrator do sistema de Rössler para $(a, b, c) = (0,36; 0,40; 4,50)$.

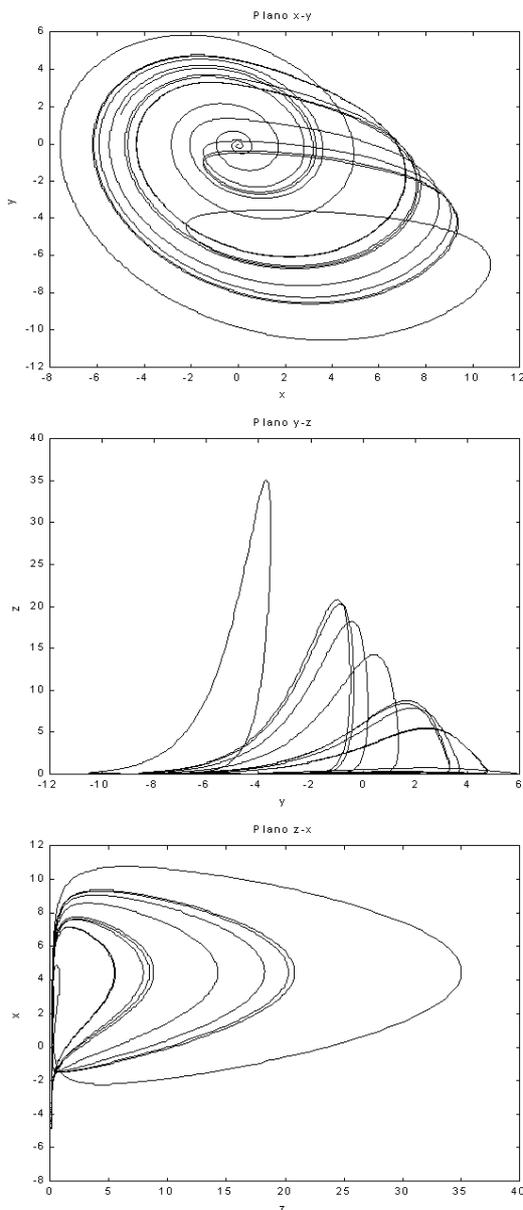


Figura 4. Comportamento dinâmico do sistema de Rössler para $(a, b, c) = (0,36; 0,40; 4,50)$.

Um conjunto de 20000 amostras é utilizado para identificação deste sistema, estes apresentados na figura 5, onde o objetivo é identificar o comportamento dinâmico da coordenada x do sistema de Rössler.

Desde que as equações (4)-(6) representam um sistema dinâmico contínuo em vez de um sistema dinâmico discreto, necessita-se utilizar um método de integração numérica. Neste caso, escolheu-se o método de Runge-Kutta de 4ª ordem com passo de 0,005 e condições iniciais nulas. O experimento foi configurado para utilizar-se 10000 amostras para a fase de estimação e 10000 amostras, para a fase de validação do modelo matemático obtido através de uma *RNWs*.

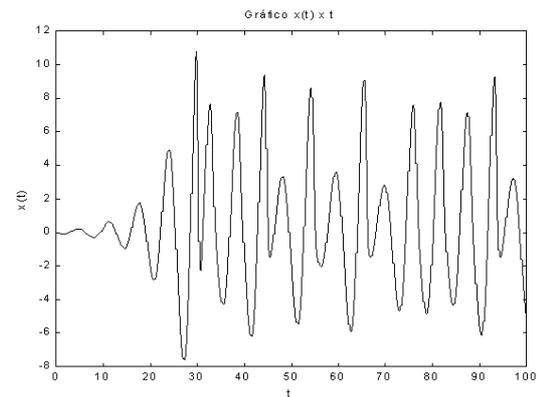


Figura 5. Série temporal da coordenada x usada para identificação do sistema de Rössler para o caso de $(a, b, c) = (0,36; 0,40; 4,50)$.

4. Resultados de simulação para os estudos de caso

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 da estrutura da *RNWs* (número de *wavelets*, escalas e espaço de busca), (iii) determinação de um procedimento de otimização da *RNWs* (adotou-se o algoritmo do gradiente estocástico [4] dos parâmetros das *wavelets* com 100 épocas de treinamento para o estudo de caso 1, e 15 épocas para o estudo de caso 2), (iv) estimação dos parâmetros do modelo matemático (*fase de estimação*), e (v) validação do modelo matemático (*fase de validação*).

O critério de desempenho do erro médio quadrático (*MSE*) é calculado pela soma do erro quadrático entre a saída real, $x(t+1)$, e a saída estimada, $\hat{x}(t+1)$, dividido pelo número de amostras consideradas do sistema dinâmico a ser identificado. Os resultados obtidos para a *RNWs* são apresentados nas tabelas 1 e 2, respectivamente.

O número inicial de *wavelets* estipulado para a camada oculta foi obtido através de um procedimento baseado no critério de Akaike [16]. O critério de Akaike mede a variância dos resíduos em relação ao número de parâmetros ajustáveis. Neste contexto, adotou-se uma abordagem de ortogonalização proposta em [17] usando

Gram-Schmidt [12] combinada com a decisão baseada no critério de Akaike [16].

Tabela 1: Resultados obtidos através da *RNWs* para o estudo de caso 1 (mapa de Hénon).

entradas da <i>RNWs</i>	wavelets na camada oculta	MSE (estimação)	MSE (validação)
$x(t-1)$	4	$3,717 \times 10^{-2}$	$4,132 \times 10^{-2}$
$x(t-1) \dots x(t-2)$	16	$6,591 \times 10^{-6}$	$5,759 \times 10^{-3}$
$x(t-1) \dots x(t-3)$	19	$1,636 \times 10^{-6}$	$5,276 \times 10^{-3}$
$x(t-1) \dots x(t-4)$	8	$1,609 \times 10^{-4}$	$1,669 \times 10^{-2}$
$x(t-1) \dots x(t-5)$	2	$1,401 \times 10^{-4}$	$3,705 \times 10^{-2}$

Tabela 2: Resultados obtidos através da *RNWs* para o estudo de caso 2 (sistema de Rössler).

entradas da <i>RNWs</i>	wavelets na camada oculta	MSE (estimação)	MSE (validação)
$x(t-1)$	2	$5,437 \times 10^{-4}$	$6,057 \times 10^{-4}$
$x(t-1) \dots x(t-2)$	29	$2,339 \times 10^{-8}$	$1,726 \times 10^{-8}$
$x(t-1) \dots x(t-3)$	35	$8,727 \times 10^{-12}$	$1,672 \times 10^{-12}$

Os resultados de identificação através de *RNWs* foram promissores para a identificação dos dois estudos de caso apresentados.

Nas figuras 6 a 9 são apresentados os gráficos relacionados ao melhor resultado (ver tabela 1) obtido com a *RNWs* para a identificação do mapa de Hénon. O melhor resultado obtido foi para uma entrada na *RNWs* com o vetor $[x(t-1) \ x(t-2) \ x(t-3)]$ (ordem 3 para um modelo auto-regressivo não-linear (*NAR*)) e usando-se 19 wavelets (também denominados de wavelons segundo a convenção adotada por [4] e [9]).

Em relação ao melhor resultado obtido, na fase de estimação, obteve-se um erro máximo de $5,337 \times 10^{-3}$, desvio padrão do erro de $-4,620 \times 10^{-3}$ e erro médio de $1,2103 \times 10^{-6}$. Entretanto, na fase de validação, quando é verificada a capacidade de generalização da *RNWs* obteve erro máximo de $1,763 \times 10^{-1}$; desvio padrão do erro de $4,023 \times 10^{-2}$ e erro médio $1,613 \times 10^{-7}$.

O erro máximo obtido pela *RNWs* foi aceitável e nota-se que pelo erro médio (tanto na fase de estimação quanto validação) que o procedimento de identificação usando *RNWs* obteve um sinal de erro decorrelacionado (ruído branco ou número com distribuição Gaussiana).

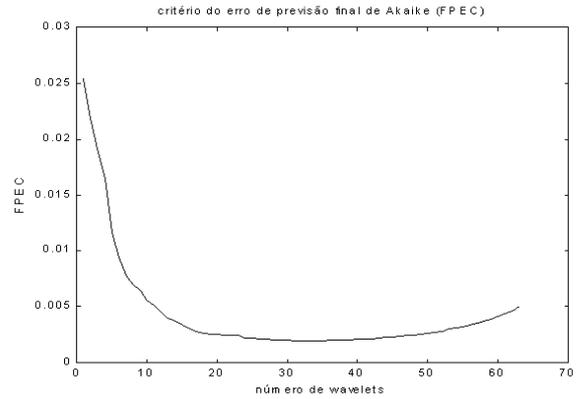


Figura 6. Resultados do critério de previsão final de Akaike para a configuração inicial da *RNWs* para identificação do mapa de Hénon (19 wavelets na camada oculta).

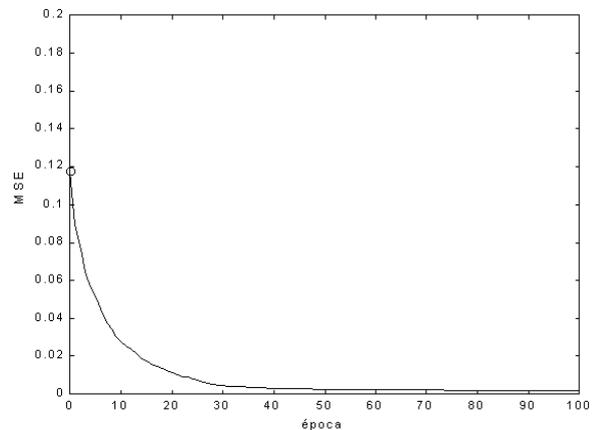


Figura 7. Convergência da previsão do mapa de Hénon usando *RNWs* (19 wavelets na camada oculta).

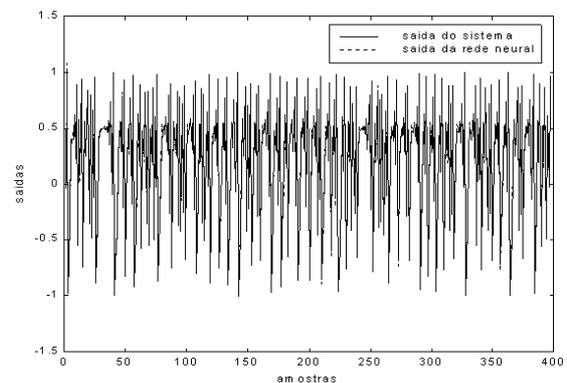


Figura 8. Resultados da previsão do mapa de Hénon (19 wavelets na camada oculta).

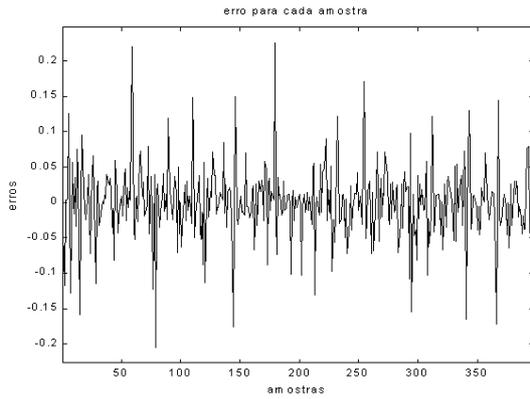


Figura 9. Erro de previsão do mapa de Hénon (19 *wavelets* na camada oculta).

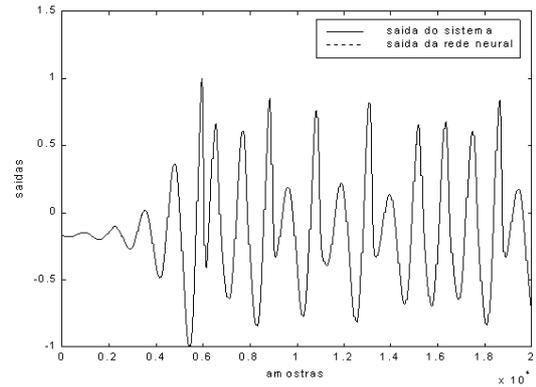


Figura 12. Resultados da previsão do sistema de Rössler (35 *wavelets* na camada oculta).

Nas figuras 10 a 13 são apresentados os melhores resultados obtidos com a *RNWs* para a identificação do sistema de Rössler, estes resumidos na tabela 2.

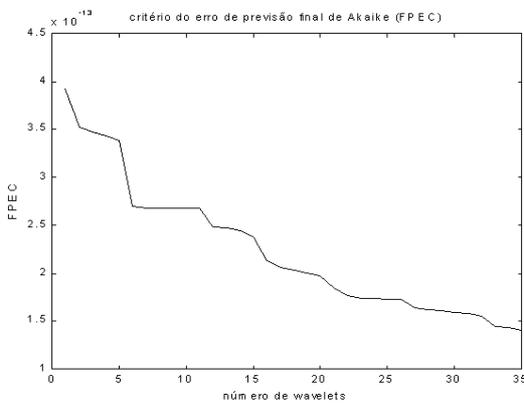


Figura 10. Resultados do critério de previsão final de Akaike para a configuração inicial da *RNWs* para identificação do sistema de Rössler (35 *wavelets* na camada oculta).

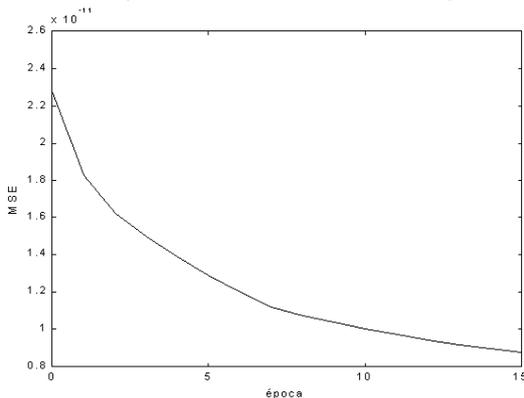


Figura 11. Convergência da previsão do sistema de Rössler usando *RNWs* (35 *wavelets* na camada oculta).

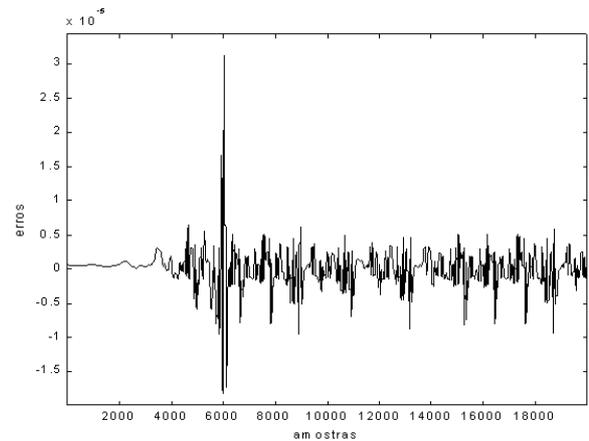


Figura 13. Erro de previsão do sistema de Rössler (35 *wavelets* na camada oculta).

O melhor resultado obtido foi para uma entrada na *RNWs* $[x(t-1) x(t-2) x(t-3)]$ com 35 *wavelets* na camada oculta da *RNWs*. Em relação ao melhor resultado preliminar obtido, na fase de estimação, obteve-se um erro máximo de $3,100 \times 10^{-6}$, desvio padrão do erro de $1,238 \times 10^{-7}$ e erro médio de $1,372 \times 10^{-7}$.

Entretanto, na fase de validação, quando é verificada a capacidade de generalização da *RNWs* foi obtido um erro máximo de $9,608 \times 10^{-6}$, desvio padrão do erro de $1,280 \times 10^{-8}$ e erro médio de $2,901 \times 10^{-8}$.

O erro máximo obtido pela *RNWs* foi aceitável e nota-se que pelo erro médio (tanto na fase de estimação quanto validação) que o procedimento de identificação usando *RNWs* obteve um sinal de erro decorrelacionado (ruído branco ou número com distribuição Gaussiana) de forma similar ao estudo de caso 1.

Em síntese, os resultados obtidos mostram a robustez da *RNWs* na previsão das séries temporais. Entretanto, alguns questionamentos podem ser realizados. A solução do problema de seleção estrutural que está presente em uma *RNWs* depende dos valores atribuídos a dilatações e translações das *wavelets*.

Neste caso, existe um compromisso, a ser estudado de forma mais aprofundada, entre o número de *wavelets*

selecionados (complexidade computacional) e a qualidade de aproximação da *RNWs* para a previsão de séries temporais com comportamento dinâmico caótico.

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 a neste artigo serem apresentados os fundamentos e a aplicação de *RNWs* para o aprendizado de sistemas com dinâmica caótica.

Foram analisados dois estudos de caso: o mapa de Hénon e o sistema de Rössler. Os resultados de simulação mostraram que as *RNWs* apresentaram aprendizado rápido, convergência rápida e boa aproximação da dinâmica caótica para os dois estudos de caso abordados. Estes resultados preliminares confirmam as propriedades das *RNWs* de identificarem sistemas dinâmicos que possuem não-linearidades e dinâmica com comportamento caótico.

Os trabalhos futuros envolverão experimentos relacionados à análise de aspectos de adaptação, robustez, atraso de transporte, correlação cruzada e algoritmos de poda para as *RNWs* quando da identificação de séries temporais com a presença de ruídos e com parâmetros variantes no tempo.

Referências

- [1] R. Brown, R. Berezdivin and L. O. Chua. Chaos and complexity, *International Journal of Bifurcation and Chaos*, vol. 11, no. 1, pp. 19-26, 2001.
- [2] G. Chen (ed.). *Controlling Chaos and Bifurcations in Engineering Systems*, CRC Press, Boca Raton, FL, 1999.
- [3] J. Wang and X. Wang. A global control of polynomial chaotic systems, *International Journal of Control*, vol. 72, no. 10, pp. 911-918, 1999.
- [4] Q. Zhang and A. Benveniste. Wavelet networks, *IEEE Transactions on Neural Networks*, vol. 3, no. 6, pp. 889-898, 1992.
- [5] H. H. Szu, X. -Y. Yang and Y. Sheng. Neural network and wavelet transform for scale-invariant data classification, *Physica Review E*, vol. 48, no. 2, pp. 1497-1501, 1993.
- [6] B. R. Bakshi and G. Stephanopoulos. Wave-net: a multiresolution, hierarchical neural network with localized learning, *AIChE Journal*, vol. 39, no. 1, pp. 57-81, 1993.
- [7] V. Kreinovich, O. Sirisaengtaksin and S. Cabrera. Wavelet neural networks are asymptotically optimal approximators for functions of one variable, *Proceedings of IEEE International Conference on Neural Networks*, Orlando, FL, USA, pages 299-304, 1994.
- [8] C. H. Chen and G. G. Lee. Multiresolution wavelet analysis based feature extraction for neural network classification, *IEEE International Conference on Neural Networks*, Washington, DC, USA, vol. 3, pages 1416-1421, 1996.
- [9] Q. Zhang. Using wavelet network in nonparametric estimation, *IEEE Transactions on Neural Networks*, vol. 8, no. 2, pp. 227-236, 1997.
- [10] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis, *IEEE Transactions on Information Theory*, 36(5): 961-1005, 1990.
- [11] M. Nikolaou and P. Vuthandam. FIR model identification: parsimony through kernel compression with wavelets, *AIChE Journal*, vol. 44, no. 1, pp. 141-150, 1998.
- [12] G. H. Golub and C. F. Van Loan. *Matrix computations*, Baltimore: The John Hopkins University Press, 1983.
- [13] M. Hénon. A two dimensional mapping with strange attractor, *Communications in Mathematical Physics*, vol. 50, pp. 69-77, 1976.
- [14] O. E. Rössler. An equation for continuous chaos, *Physical Letters*, vol. 35A, pp. 397-398, 1976.
- [15] K. T. Alligood, T. D. Sauer and J. A. Yorke. *Chaos: an introduction to dynamical systems*, Springer, 1996.
- [16] R. Johansson. *System modeling identification*. Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.
- [17] J. R. Hull and H. P. Pendse. A neural network algorithm using wavelets and auto regressive inputs for system identification, *International Conference on Neural Networks*, Houston, TX, USA, vol. 2, pp. 728-732.