

Validação de Modelos Neurais Identificados a Partir de Um Sistema Caótico

Marcelo V. Corrêa^{1†}, Luis A. Aguirre^{2†}, Antônio de P. Braga^{3‡},

¹Instituto Católico de Minas Gerais – ICMG
Av. Tancredo Neves 3500, 35170-056 Cel Fabriciano, MG, Brazil

[†] Laboratório de Modelagem, Análise e Controle de Sistemas Não Lineares

[‡] Laboratório de Inteligência e Tecnologia Computacional

^{2,3} Departamento de Engenharia Eletrônica Universidade Federal de Minas Gerais
Av. Antônio Carlos 6627, 31270-901 Belo Horizonte, M.G., Brazil

mcorrea@gold.com.br, aguirre@cpdee.ufmg.br, apbraga@cpdee.ufmg.br

Abstract

The present work describes some results obtained when modeling a chaotic system using Multilayer Perceptrons. Such a system is considered as a benchmark for problems of identification of non linear systems because it presents a great diversity of dynamic regimes including chaos. The results described here show that several nets that satisfied the same stop criterion have very different dynamic behaviors. Besides, such behaviors become evident when techniques such as Poincaré maps and bifurcation diagrams are used to verify Neural Network generalization. The results suggest that when the objective of the training of a Neural Network is to model non linear dynamic systems, other approaches, besides the minimization of the prediction error should be used in model validation.

1. Introdução

Um dos grande desafios em ciência e tecnologia é a identificação de sistemas, ou seja, a obtenção de modelos matemáticos a partir de dados observados, sem conhecimento prévio. Uma das questões básicas na identificação de sistemas é a escolha da representação a ser utilizada. Na área de modelos lineares, uma das representações mais utilizadas é o modelo ARMAX [1]. Recentemente, grande atenção tem sido dada às representações não lineares. Uma lista completa de tais representações seria longa demais para ser incluída aqui, mas algumas daquelas que têm recebido significativa atenção na literatura incluem modelos polinomiais erracionais contínuos [2], discretos (NARMAX) [3], funções de base radial [4, 8] e Redes Neurais Artificiais [5, 6, 7].

O presente trabalho descreve alguns resultados

obtidos ao modelar um sistema caótico utilizando Multilayer Perceptron [11, 12]. O sistema caótico se presta como benchmark para problemas de identificação de sistemas não lineares por apresentar uma grande diversidade de regimes dinâmicos que incluem regimes caóticos. Os resultados aqui descritos mostram que diversas redes que foram validadas pela predição de k-passos à frente têm comportamentos dinâmicos muito distintos. Além disso, tais comportamentos tornam-se evidentes na medida em que técnicas do tipo mapas de Poincaré e diagramas de bifurcação são usados para verificar o poder de generalização das redes. Os resultados sugerem que quando o objetivo do treinamento de uma RNA for modelar a dinâmica não linear do sistema original, outros critérios, além da predição de k-passos devem ser usados na validação dos modelos.

A validação dos modelos, que tem relevância significativa em qualquer problema de identificação, adquire uma maior importância no caso de sistemas que exibem diversos tipos de regimes dinâmicos incluindo o caos [9, 6]. Dentre as possibilidades de validação destaca-se a reconstrução de atratores, os diagramas de bifurcação e a obtenção de seções de Poincaré. Uma das principais limitações de se utilizar tais ferramentas em problemas práticos é que nem sempre é possível obter tais diagramas para os sistemas originais. Sendo assim, fica impossível comparar os diagramas de bifurcação e as seções de Poincaré dos modelos com aqueles do sistema original. Por outro lado, em estudos onde o sistema original é um modelo do qual tais diagramas podem ser facilmente obtidos, o seu uso é recomendado para investigar mais a fundo as características dinâmicas dos modelos obtidos. É fundamental perceber que na maioria dos casos usar diagramas de bifurcação e seções de Poincaré para avaliar o desempenho dinâmico dos modelos em muito excede ao procedi-

mento convencional de verificar como tais modelos predizem os dados. Além disso, a variância (ou valor absoluto) de erros de predição de um passo a frente, tão comumente usada como critério de parada de treinamento, é totalmente ineficaz na caracterização do desempenho de modelos dinâmicos. Finalmente, menciona-se que a discussão acima é válida para representações de modelos não lineares em geral, e não apenas para redes neurais [10]

Apresenta-se aqui um estudo sobre a capacidade que algumas topologias de redes neurais artificiais (RNA), têm de reproduzir a dinâmica caótica observada na equação de Duffing-Ueda [13, 14].

2 Sistema a ser modelado

O exemplo de sistema não linear utilizado neste trabalho é a equação de Duffing-Ueda. Esta equação foi originalmente proposta como um modelo de um oscilador não linear e tem sido considerada como um padrão no estudo de dinâmicas não lineares. O modelo é simples e pode produzir uma grande variedade de regimes dinâmicos [13].

A equação de Duffing-Ueda é escrita como [14]

$$\ddot{y} + k\dot{y} + y^3 = u(t), \quad (1)$$

$k = 0,1$ foi usado para produzir os dados originais via simulação digital, estes dados foram os mesmos usados por [9]. Um algoritmo de Runge-Kutta de quarta-ordem com um intervalo de integração de $\pi/3000s$ foi usado para simular a resposta do sistema para uma determinada entrada. Entrada $u(t)$ e saída $y(t)$ foram uniformemente amostrados com período de $T_s = \pi/60s$, correspondendo a 120 pontos por pseudo período do sinal de entrada. A massa de dados usada para identificação está apresentada na Figura 1. Todas as Figuras usadas neste trabalho que se referem ao sistema original foram retiradas de [9].

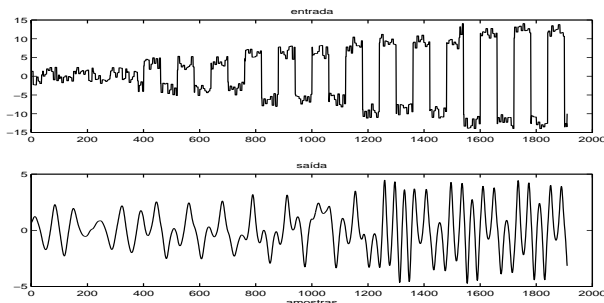


Figura 1: Dados do sistema Duffing-Ueda usados para identificação

3 Modelos Identificados

Os modelos apresentados neste trabalho, foram obtidos utilizando as rotinas do “toolbox” de identificação baseados em redes neurais, para o MATLABTM desenvolvido por M. Norgaard [11]. A rotina usada foi a `nnarmax1.m`, que identifica uma rede neural ARMAX (auto-regressive, moving average with exogenous input) a partir de um par de vetores contendo entrada e saída desejada. A rede é obtida a partir da escolha de: n_y , n_u e n_e , respectivamente, número de atrasos da saída, entrada e número de termos de ruído, além da seleção do número de neurônios na camada escondida e da funções de ativação usadas.

Como critério de parada para o algoritmo de treinamento, utilizaram-se sempre valores default da rotina. Os valores de n_y , n_u e n_e usados foram, respectivamente, 3, 2 e 10. Os valores de n_y e n_u foram escolhidos iguais aos “lag’s” usados por [9]. Empiricamente, escolheu-se $n_e = 10$.

Obtiveram-se várias redes variando apenas o número de neurônios na camada escondida. As redes obtidas foram chamadas: *nn2*, *nn3*, *nn4*, *nn5*, *nn6*, *nn7*, *nn8*, *nn9*, *nn10*, *nn11*, *nn12*, onde os números 2, 3, ... 12 correspondem ao número de neurônios na camada escondida. Em todos os casos utilizou-se a tangente hiperbólica para função de ativação na camada escondida, e função de ativação linear para a camada de saída. Os pesos foram sempre inicializados com valores nulos.

Para todas as topologias usadas na obtenção dos modelos, utilizou-se o mesmo critério de parada. Embora tenha ocorrido número de épocas diferentes para cada topologia usada, sempre o critério de parada foi atingido.

4 Validação das redes obtidas

A questão crítica na obtenção de modelos é a sua validação, principalmente se o sistema possui dinâmica caótica [9]. Para o sistema apresentado neste trabalho, utilizaram-se quatro critérios de validação:

- Predição da série temporal
- Reconstrução do atrator
- Seção de Poincaré
- Diagramas de Bifurcação

Detalhes sobre cada um dos critérios de validação referidos acima podem ser encontrados em [15, 16].

4.1 Predição da série temporal

Consiste em simular o modelo com a mesma entrada do sistema, e verificar sua capacidade de seguir

os dados originais. Neste caso a predição é feita k passos a frente. Nesta simulação utilizou-se a mesma massa de dados usada para a identificação. Os resultados estão apresentados na Figura 2.

Na predição de k -passos à frente não foi possível perceber diferenças significativas no comportamento dos modelos, principalmente considerando o comportamento caótico observado nos dados de identificação. A razão para isto é o fato de que sistemas caóticos têm pelo menos um expoente de Lyapunov positivos. A consequência disso é que no espaço de estados, numa “direção” as trajetórias divergem. Em função disso, pequenas diferenças entre condições iniciais são amplificadas à medida que o sistema se desenvolve. Isso é conhecido como sensibilidade a condições iniciais e é um limite teórico e prático à predição (de médio e longo prazo) de séries temporais caóticas. A Figura 2 apresenta um resultado típico obtido na simulação das redes.

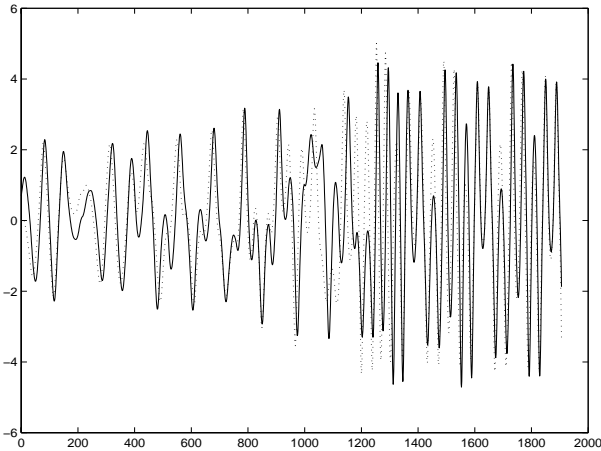


Figura 2: Resultado típico obtido para todas as redes, na predição de k -passos a frente. (-) Sistema Original e (..) Rede

4.2 Reconstrução do atrator

Neste critério é usado um pseudo-plano de fases constituído por $y(t)$ e $y(t - T_p)$ para reconstruir o atrator [17, 18].

As trajetórias obtidas, a partir da simulação dos modelos identificados, estão apresentadas nas Figuras 3(a)-(f). Os modelos foram simulados com a entrada $u(t) = A \cos(\omega t)$, com $A = 11$ e $\omega = 1$ rad/s. A Figura 3(f) apresenta a trajetória obtida a partir do sistema original com $T_p = 200 \times \pi/3000$, sendo que para os modelos utilizou-se $T_p = 4 \times \pi/60$, equivalendo sempre ao tempo de quatro amostras.

Os atratores reconstruídos com modelos com 4, 5, 6, 7, 8, 10 e 11 neurônios na camada escondida estão aparentemente próximos do atrator do sistema original. A fim de ser categóricos nesta afirmação, não bastaria olhar apenas para a geometria das projeções

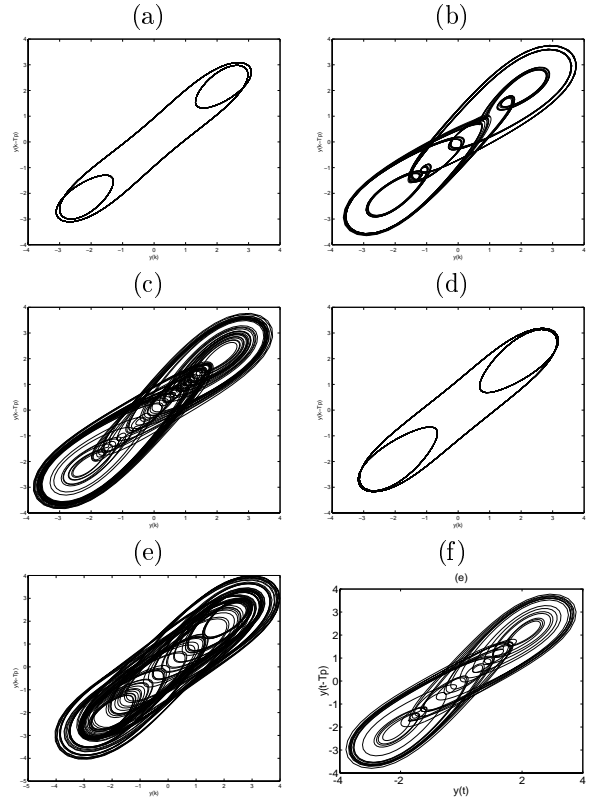


Figura 3: Trajetória no pseudo-plano de fases (a) Modelo $nn2$, (b) Modelo $nn3$, (c) Modelo $nn8$, (d) Modelo $nn9$, (e) Modelo $nn12$, (f) Sistema original. Eixos: (a) a (d) e (f) $\rightarrow -4$ a 4 ; (e) $\rightarrow -5$ a 4

dos atratores, mas seria necessário verificar como as trajetórias se cruzam no espaço de estados. A análise topológica inclui esse tipo de investigação [19]. As redes, com dois, três e nove neurônios na camada escondida, apresentaram comportamento periódico, e a rede com 12 neurônios apresentou uma deformação do atrator, com um deslocamento para a esquerda.

4.3 Seção de Poincaré

Mapa de primeiro retorno de um sistema dinâmico de ordem n , ou mapa de Poincaré, é uma hipersuperfície de dimensão $n-1$ que é transversal ao fluxo $\phi \in \mathbf{R}^n$. O mapa de Poincaré $P = \Sigma \rightarrow \Sigma$ é definido para um ponto $q \in \Sigma$, por:

$$P(q) = \phi_\tau(q), \quad (2)$$

onde τ é o tempo para a órbita $\phi_t(q)$ baseado em q , retornar pela primeira vez em Σ .

As Figuras 4(a)-(f) apresenta as seções de Poincaré obtidas a partir da simulação de alguns modelos para entrada, $u = 11 \cos(t)$, e a Figura 4(f) apresenta a seção de Poincaré para o sistema original. As redes omitidas apresentaram comportamentos similares, mas certamente não idênticos, à Figura 4(c).

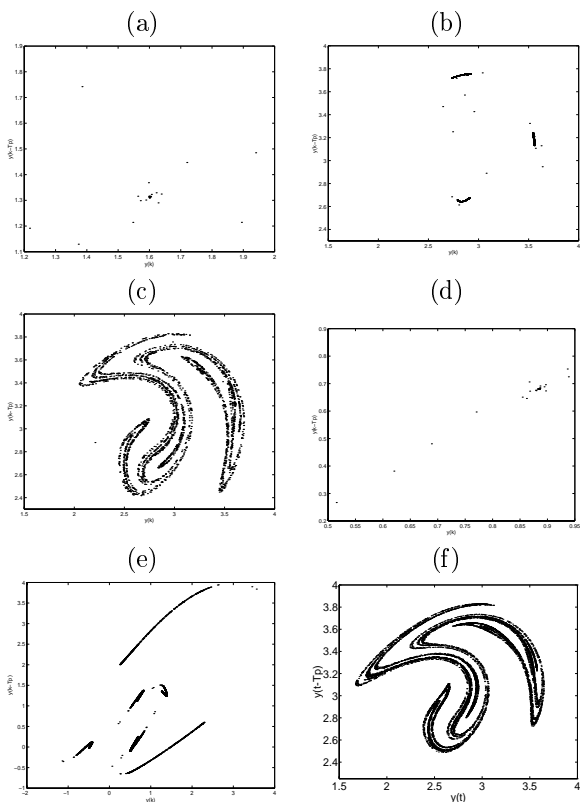


Figura 4: Seção de Poincaré para a entrada $u = 11\cos(t)$ (a) Modelo $nn2$, período-1, (b) Modelo $nn3$, período-3, (c) Modelo $nn8$, caos, (d) Modelo $nn9$, período-1, (e) Modelo $nn12$, possivelmente quasi-periódico, e (f) Sistema original, caos. Eixos: (a) $\rightarrow 1,2$ a 2 ; (b),(c) e (f) $\rightarrow 1,5$ a 4 ; (d) $\rightarrow 0,5$ a $0,9$; (e) $\rightarrow -2$ a 4

Segundo este critério as redes que mais se aproximaram do comportamento do sistema original foram as redes com 4, 5, 6, 7, 8, 10 e 11 neurônios na camada escondida. Os modelos com 2, 3, 9 e 12 neurônios possuem seções de Ponincaré muito distantes do sistema original.

4.4 Diagrama de Bifurcação

Um ponto r de um diagrama de bifurcação, para o caso considerado neste trabalho, é definido como:

$$r = \left\{ \begin{array}{l} (y, A) \in \mathbf{R} \times I | y = y(t_i), \\ A = A_0; t_i = K_{ss} \times 2\pi/w \end{array} \right\} \quad (3)$$

onde I é o intervalo $I = [A_i A_f] \subset \mathbf{R}$. O ponto r é obtido por simulação do sistema por um tempo suficientemente longo para atingir o regime permanente, então é feito o gráfico de $y(K_{ss} \times 2\pi)$ por A_0 . Para cada valor de A , tem-se n_b pontos.

O diagrama de bifurcação dá uma idéia dos valores de $A \in I$ para os quais o sistema bifurca e qual é o tipo de bifurcação. No estudo de sistemas não lineares o diagrama é usado para determinar qual é

a faixa de valores do parâmetro de controle em que o sistema possui comportamento caótico. Deve ser notado que ao passo que o mapa de Poincaré fornece uma imagem (um corte) do atrator no espaço de estados, o diagrama de bifurcação resume os diversos regimes dinâmicos exibidos pelo sistema para diversos valores do parâmetro de bifurcação.

Os diagramas de bifurcação para os modelos identificados estão apresentados nas Figuras 5(a)-(e). O parâmetro de bifurcação usado foi a amplitude do sinal senoidal da entrada, na faixa $4,5 < A < 12$. 150 pontos equidistantes foram gerados com $K_{ss} = 63$ e $n_b = 20$. O diagrama de bifurcação do sistema original está apresentado na Figura 5(f), tendo sido gerado com $K_{ss} = 400$ e $n_b = 20$ [9]. A diferença nos valores de K_{ss} e o número de pontos usados, para o modelo e para o sistema original se deve ao grande esforço computacional necessário para simular os modelos baseados em rede neurais. $K_{ss} = 400$ significa simular a rede com 48000 pontos, 750 vezes, o que que acredita-se não ser necessário, pois $K_{ss} = 63$, já é suficiente para que o sistema atinja o regime permanente.

É interessante notar que o mapa de Poincaré corresponde a um valor no diagrama de bifurcação. Por exemplo, todos os mapas da Figura 4 correspondem aos atratores observados para $A = 11$, nos respectivos diagramas da Figura 5. Portanto, os regimes dinâmicos (para $A = 11$) “período-1” e caos podem ser constatados com certa facilidade. Os regimes “período-3” e quasi-periódicos não são tão facilmente discerníveis nos diagramas de bifurcação.

Segundo este critério de validação, verifica-se que as redes que mais se aproximaram do comportamento do sistema original foram as redes com 4, 5, 6, 7, 8, 10 e 11 neurônios na camada escondida. Porém resalta-se o fato de nenhuma delas reproduzirem a seqüência de bifurcações observadas nos dados originais.

5 Comentários

Os resultados apresentados na seção anterior apontam para a necessidade de se estabelecer critérios de validação que não sejam baseados apenas na predição de k -passos, quando se trata de modelagem de sistemas não lineares. Verifica-se que apenas o critério de parada do algoritmo de treinamento de uma rede neural não é suficiente para garantir comportamento dinâmico da rede igual ao observado no sistema original. Isso não é uma propriedade peculiar a redes neurais, pelo contrário, praticamente todas as representações não lineares sofrem desses problemas. O que diferencia uma representação de outra é a facilidade com que tais problemas podem ser resolvidos em casos práticos.

Por outro lado, a reconstrução do atrator, conforme visto na Figura 3 também não foi suficiente para garantir a capacidade de generalização das redes

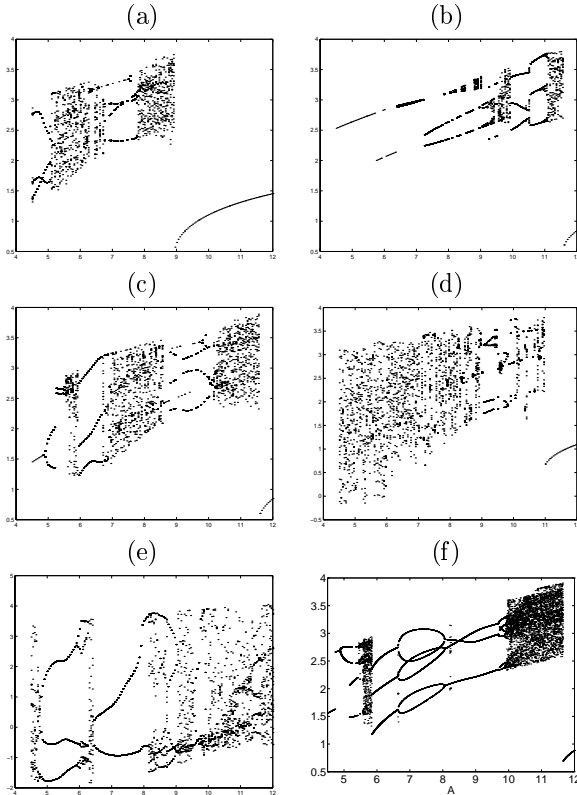


Figura 5: Diagramas de bifurcações (a) Modelo *nn2*, (b) Modelo *nn3*, (c) Modelo *nn8*, (d) Modelo *nn9*, (e) Modelo *nn12* e (f) Sistema original. Eixos: Abscissas (todas) 4,5 a 12, ordenadas: (a),(b),(c) e (f) \rightarrow 0,5 a 4; (d) \rightarrow -0,5 a 4; (e) \rightarrow -2 a 5

obtidas, apesar de que uma análise topológica [19] seria necessária a fim de quantificar quão próximos estão os atratores do atrator original. Observa-se também que a seção de Poincaré se mostrou mais apropriada para estabelecer comparações entre as redes obtidas, para $A = 11$.

Os diagramas de bifurcação, apresentam como uma ferramenta mais poderosa para se verificar a capacidade de generalização das redes neurais. Observa-se que, nos casos apresentados aqui, nenhuma rede conseguiu reproduzir a seqüência de bifurcações observadas nos dados originais. Neste ponto cabe um estudo mais aprofundado sobre a influência do número de camadas escondidas, e do número de nodos, assim como da influência das funções de ativação usadas.

Antes de encerrar esta seção deve ser notado que existem modelos identificados a partir dos dados da Figura 1 que reproduzem satisfatoriamente os critérios aqui usados bem como outros tais como o máximo expoente de Lyapunov e a dimensão de correlação [9].

6 Conclusão

Os resultados apresentados neste trabalho mostram que apenas a análise de erros de predição não é suficiente para validar modelos de sistemas que exibem diversos regimes dinâmicos inclusive caos. Várias redes que atingiram o critério de parada do algoritmo de treinamento, e passaram pela validação de k -passos, não reproduziram o comportamento observado no sistema dinâmico original.

Observa-se que a capacidade de generalização de uma rede neural usada na identificação de sistemas não lineares deve ser testada através de diagramas de Poincaré, e diagramas de Bifurcação. Redes que reproduziram o atrator para uma determinada entrada, não conseguiram reproduzir nenhuma das seqüências de bifurcação observadas nos dados originais, bem como não reproduziram a seção de Poincaré observada no sistema original.

Agradecimentos

Agradecimentos ao financiamento à pesquisa do CNPq e PRPq/UFMG. A pesquisa do MVC é financiada pelo ICMG.

Referências

- [1] L. Ljung. System identification: Theory for the user. Prentice Hall, New Jersey, 1987.
- [2] G. Gouesbet and C. Letellier. Global vector field reconstruction by using a multivariate polynomial L_2 approximation on nets. Phys. Rev. E, 49(6):4955–4972, 1994.
- [3] S. Chen and S. A. Billings. Representation of non-linear systems: the NARMAX model. Int. J. Control, 50(5):1873–1896, 1989.
- [4] M. Casdagli. Nonlinear prediction of chaotic time series. Physica D, 35:335–356, 1989.
- [5] K. S. Narendra and K. Parthasarathy. Identification and control of dynamical systems using neural networks. IEEE Trans. Neural Net., 1:4–27, 1990.
- [6] H. M. Henrique, E. L. Lima, and J. C. Pinto. A bifurcation study on neural network models for nonlinear dynamics systems. Latin American Applied Research, 28:187–200, 1998.
- [7] R. Gençay and T. Liu. Nonlinear modelling and prediction with feedforward and recurrent networks. Physica D, (108):119–134, 1997.
- [8] Q. M. Zhu and S. A. Billings. Fast orthogonal identification of nonlinear stochastic models and radial basis functions. Int. J. Control, 64(5):871–886, 1996.

- [9] L. A. Aguirre and S. A. Billings. Validating identified nonlinear models with chaotic dynamics. *Int. J. Bifurcation and Chaos*, 4(1):109–125, 1994.
- [10] L. A. Aguirre. Some remarks on structure selection for nonlinear models. *Int. J. Bifurcation and Chaos*, 4(6):1707–1714, 1994.
- [11] M. Norgaard. *Neural Network Based System Identification - TOOLBOX*. DTU, Technical University of Denmark, 1997. Technical Report 97-E-851.
- [12] A. P. Braga, A. P. L. T. Carvalho, and T. B. Ludermit. *Fundamentos de Redes Neurais Artificiais*. 11^a Escola de Computação, Rio de Janeiro, Brasil, 1998.
- [13] Y. Ueda. Steady motions exhibited by Duffing's equation: A picture book of regular and chaotic motions. In P. J. Holmes, editor, *New approaches to nonlinear problems in dynamics*, pages 311–322. SIAM, 1980.
- [14] Y. Ueda. Random phenomena resulting from nonlinearity in the system described by Duffing's equation. *Int. J. Non-Linear Mech.*, 20(5/6):481–491, 1985.
- [15] L. Aguirre. A tutorial introduction to nonlinear dynamics and chaos, part i: Tools and benchmarks. *Controle & Automação*, 1(7):29–49, 1996.
- [16] L. Aguirre. A tutorial introduction to nonlinear dynamics and chaos, part ii: Modeling and control. *Controle & Automação*, 1(7):50–66, 1996.
- [17] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw. Geometry from a time series. *Phys. Rev. Lett.*, 45(9):712–716, 1980.
- [18] F. Takens. Detecting strange attractors in turbulence. In D. A. Rand and L. S. Young, editors, *Dynamical systems and turbulence*, Lecture Notes in Mathematics, vol. 898, pages 366–381. Springer Verlag, Berlin, 1980.
- [19] R. Gilmore. Topological analysis of chaotic dynamical systems. *Rev. Modern Phys.*, 70(4):1455–1529, 1980.