Physics Letters A 380 (2016) 696–706

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

NARX prediction of some rare chaotic flows: Recurrent fuzzy functions approach

Sobhan Goudarzi [a], Sajad Jafari [a,*], Mohammad Hassan Moradi [a], J.C. Sprott [b]

[a] Biomedical Engineering Department, Amirkabir University of Technology, Tehran 15875-4413, Iran
[b] Department of Physics, University of Wisconsin–Madison, Madison, WI 53706, USA

ARTICLE INFO

Article history:
Received 24 September 2015
Received in revised form 26 November 2015
Accepted 27 November 2015
Available online 30 November 2015
Communicated by C.R. Doering

Keywords:
Rare chaotic flows
Hidden attractor
Fuzzy functions
Recurrent structure
Long term prediction

ABSTRACT

The nonlinear and dynamic accommodating capability of time domain models makes them a useful representation of chaotic time series for analysis, modeling and prediction. This paper is devoted to the modeling and prediction of chaotic time series with hidden attractors using a nonlinear autoregressive model with exogenous inputs (NARX) based on a novel recurrent fuzzy functions (RFFs) approach. Case studies of recently introduced chaotic systems with hidden attractors plus classical chaotic systems demonstrate that the proposed modeling methodology exhibits better prediction performance from different viewpoints (short term and long term) compared to some other existing methods.

1. Introduction

Most familiar examples of low-dimensional chaotic flows occur in systems having one or more saddle points. Such saddle points allow homoclinic and heteroclinic orbits and the prospect of rigorously proving the chaos when the Shilnikov condition is satisfied. Furthermore, such saddle points provide a means for locating any strange attractors by choosing an initial condition on the unstable manifold in the vicinity of the saddle point. Such attractors have been called "self-excited," and they are overwhelmingly the most common type described in the literature.

Recently, many new chaotic flows have been discovered that are not associated with a saddle point, including ones without any equilibrium points, with only stable equilibria, or with a line containing infinitely many equilibrium points [1–18]. The attractors for such systems have been called "hidden attractors" [19–30], and that accounts for the difficulty of discovering them since there is no systematic way to choose initial conditions except by extensive numerical search. Hidden attractors are important in engineering applications because they allow unexpected and potentially disastrous responses to perturbations in a structure like a bridge or aircraft wing.

In this work we propose a new method for predicting the global behavior of chaotic flows with hidden attractors. It is known that the long-term prediction of chaotic time series is not possible due to the sensitive dependence on initial conditions [31] and that their prediction is much more difficult than for static/algebraic systems [32]. However it is still useful to find a model which can provide short-term prediction or can reproduce the geometrical properties of a chaotic system, such as the shape of the strange attractor.

Different approaches have been used for chaotic signal prediction. Fuzzy Function (FF) systems represent one of the recent interesting soft computing approaches used in various applications such as modeling, classification, and prediction [33]. Turksen introduced this type of fuzzy structure [33–35] which is simpler compared to neuro-fuzzy rule-based systems. The multidimensional input space of FFs leads to an elimination problem due to the projection onto each axis. This is one of the main differences between multidimensional structures and rule-based structures [36]. Consequently, the obtained membership values besides the input variables are used to estimate fuzzy functions. Different regression methods like Least Square Estimation (LSE) [35], Multi Adaptive Regression Spline (MARS) [37], and Support Vector Machine (SVM) [38] can be used to estimate these functions.

With the addition of recurrent structures to a model responding to memory information based on prior system states, a significant increase in addressing the temporal sequence capability can be achieved [39–44]. In this way, some literature exists on the combi-

* Corresponding author.
E-mail address: sajadjafari@aut.ac.ir (S. Jafari).

http://dx.doi.org/10.1016/j.physleta.2015.11.036

nation of recurrent structures and fuzzy systems in two categories of local and global feedback. Juang et al. [41] employed local-rule feedback and took advantage of a variable-dimension Kalman filter for learning. Lin et al. [40] introduced recurrent self-evolving neuro-fuzzy networks that have a local and global link in the aggregation step. Ganjefar and Toghi placed a single neuron with a mother wavelet activation function and local feedback in each rule to achieve better results by modifying the learning algorithm especially in on-line applications [42]. Tellez et al. based on passivity theory and by the use of an online recurrent layer, introduced inverse optimal controllers which are trained by an extended Kalman filter [45]. Theocharis implemented recurrent structure with a context node in the form of an FIR synaptic filter and achieved enhanced temporal capacity in a higher-order system for modeling a complex nonlinear temporal process [39]. This paper proposes Recurrent Fuzzy Functions (RFFs) which have the following salient characteristics:

(1) A novel FF structure that benefits both interactive rules and recurrent structures is proposed. Before this work, FF systems were used with weighted averages based on rule firing rates for aggregation, but in this study interactively recurrent nodes are used to improve the learning capacity of the dynamical structures of a time series.

(2) One important task in designing recurrent systems is training of the feedback weights. In our proposed system nodes, weights are trained with steepest descent that automatically tune the learning rate with a line search based on the strong Wolf condition because of its fast learning speed.

(3) The computation is more efficient, and structure is simpler than other considered fuzzy structures, and it is more generalizable.

Also by use of RFFs as the nonlinear autoregressive with exogenous input (NARX) model of the data, prediction of chaotic flows with a hidden attractor is investigated using two different strategies: short-term (quantitative) and long-term (qualitative).

The rest of this paper is organized as follows: the preliminaries of the NARX model are briefly reviewed in section 2. Then in section 3, the FCM method for clustering will be presented. MARS regression is described in section 4. Details of the proposed RFFs' structure and parameter learning scheme are presented in section 5. In section 6, we introduce some rare chaotic systems with a hidden attractor which will be examined in case studies. Sections 7 and 8 give results and conclusions.

2. NARX model and optimal parameter

Generally in statistical prediction, a stochastic model based on previous observation is constructed to predict current and future values. A popular type of such a model is the nonlinear autoregressive moving average model with exogenous inputs (NARMAX) which is given by [46]:

$$y(t) = F\left[\, y(t-1), \ldots, y(t-n_y), e(t-1), \ldots, e(t-n_e), x(t-1), \ldots, x(t-n_x)\right] + e(t) \tag{1}$$

where x, e and y are external input, noise (which can be seen as representing the prediction error), and output of the system, respectively. F is an unknown nonlinear function, and n_x, n_e and n_y are the maximum lags of the input, noise, and output, respectively. A special case of the general model is the NARX (n_y, n_x) model:

$$y(t) = F\left[\, y(t-1), \ldots, y(t-n_y), x(t-1), \ldots, x(t-n_x)\right] + e(t) \tag{2}$$

where it is assumed that $e(t)$ has zero mean and finite variance σ^2 and is independent and identically distributed. The predictor model in many problems can be designed without the use of external input. Here we just use the past time series values and the prediction error.

The method of making a NARX representation involves determining the structure and estimating the parameters of the unknown nonlinear system from data. Here we use the proposed RFFs as the structure, and parameters are estimated that minimize the prediction error.

3. Fuzzy C-means clustering

The cost function of the basic FCM algorithm assuming a known number of clusters is as follows [47]:

$$J_q(V,U) = \sum_{i=1}^{M}\sum_{j=1}^{C} u_{ij}^{q} d(x_i, v_j) \tag{3}$$

subject to the constraints:

$$\sum_{j=1}^{C} u_{ij} = 1, \quad i = 1, \ldots, M \tag{4}$$

where

$$u_{ij} \in [0,1], \qquad 0 < \sum_{i=1}^{M} u_{ij} < M \quad i = 1, \ldots, M,\ j = 1, \ldots, C$$

and $q > 1$ is the fuzziness, and x_i, v_j, M, and C are the ith input data, the center of the jth cluster, the number of data points, and the number of clusters, respectively. Also U is an $M \times m$ matrix whose ijth element is the membership degree of x_i in the jth cluster, and V is a $C \times m$ matrix which contains the m-dimensional centers of the clusters, and $d(x_i, v_j)$ is the distance between x_i and the jth cluster center.

After minimization, a closed form for the degree of membership of the features in the clusters is as follows [47]:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{d(x_i, v_j)}{d(x_i, v_k)}\right)^{\frac{1}{q-1}}} \tag{5}$$

and the cluster prototype is:

$$v_j = \frac{\sum_{i=1}^{M} u_{ij}^{q} x_i}{\sum_{i=1}^{M} u_{ij}^{q}} \tag{6}$$

Fuzzy partition is carried out through an iterative process consisting of computing the degree of membership and center of the clusters by use of Eqs. (5) and (6), respectively, with random initialization. Based on the global convergence theorem of Zangwill [48], when different distance measures that satisfy certain conditions discussed in [49] are employed, convergence of the sequence produced by the above algorithm in a finite number of iterations to a local minima, has been proved [50].

4. Multi adaptive regression spline

MARS is one of the adaptive regression methods. This nonparametric regression approach can be considered a generalization of stepwise linear regression and can efficiently represent the nonlinear relation and hidden patterns in data sets [51]. The sum of squares error for a general regression is as follows:

$$SSE = \sum_{i=1}^{n}\left(y_i - \beta_0 - \beta_1 b_1(x_i) - \cdots - \beta_P b_P(x_i)\right)^2 \tag{7}$$