GENETIC PROGRAMMING FOR DYNAMIC CHAOTIC SYSTEMS MODELLING

Katya Rodríguez-Vázquez
ACSE, University of Sheffield
Mappin Street S1 3JD, U.K.
katya@acse.shef.ac.uk

Peter J. Fleming
ACSE, University of Sheffield
Mappin Street S1 3JD, U.K.
P. Fleming@sheffield.ac.uk

Abstract- This work presents an investigation into the use of Genetic Programming (GP) applied to chaotic systems modelling. A difference equation model representation was proposed for being the basis of the hierarchical tree encoding in GP. Based upon the NARMA difference equation model and formulating the identification as a multiobjective optimisation problem, Chua's circuit was studied.

The formulation of the GP fitness function, defined as a multiobjective function, generated a set of non-dominated chaotic models. This approach considered criteria related to the complexity, performance and also statistical validation of the models in the fitness evaluation. The final set of non-dominated model solutions were able to capture the dynamic characteristics of the system and reproduce the chaotic motion of the double scroll attractor.

1 INTRODUCTION

Genetic programming (GP) is an evolutionary paradigm where the computer structures which undergo adaptation are themselves represented as computer programs (Koza, 1992). Since the emergence of the GP paradigm, there have been an increasing number of researchers working on both theory and applications of genetic programming. One of these application areas is time series prediction. This problem is formulated as the discovery of programs that produces a near-optimal model of the system under investigation which can reproduce its original behaviour.

Time series analysis and prediction are important in the study of a wide class of signal processing problems where applications range from the analysis of marketing data, brain wave patterns, and signals from the vibration testing of mechanical structures through to speech processing. In each case a model considered the best possible approximation to the dynamic system is estimated, based on observation data.

However, it has long been realised that the responses of many non-linear dynamic systems do not follow simple, regular, and predictable trajectories, but swirl around in a random-like and seemingly irregular behaviour. As long as the process involved is non-linear, even a simple strictly deterministic model may develop such complex behaviour. This behaviour is known and defined as ***chaos*** and has led to developments in the study of non-linear systems.

A special feature of chaos is its fundamental property, known as *extreme sensitivity* (of the system dynamics) to initial conditions, in the sense that two sets of similar initial conditions can give rise to two dramatically different asymptotic states of the system trajectory. A second feature of chaos is the inability to predict long-term behaviour.

Some studies of GP used for chaotic time series predictionhave been undertaken. Koza (1992) introduced the first studies on the discovery of programs for fitting time series data. His study was based upon the short term prediction of the logistic equation. Although this example is one of the simplest equations that exhibits chaotic motion, it was the starting point for further investigations.

Mulloy *et al.* (1996) have also explored the area of chaotic prediction by using GP. They have presented a comparison of previous results by Oakley (1994) and Iba *et al.* (1993) based upon the Mackey-Glass equation. Iba *et al.* (1993) introduced a hybrid identification system called STROGANOFF (Structured Representation On Genetic Algorithms for Non-linear Function Fitting). This system was a combination of both GP and a method of heuristic self-organisation GMDH (Ivakhnenko, 1971). However, the prediction task was short-term instead of long-term as was the study carried out by Oakley.

The approach of Koza, Oakley and Mulloy *et al.* had the same foundations, where the function set was defined as

$$\mathbf{F} = \{+, -, *, \%, \sin, \cos, \exp, \log\}$$

An alternative chaotic non-linear system identification (prediction) approach is described in this paper. This approach uses GP for evolving potential models that can reproduce the dynamic behaviour exhibited by chaotic systems. The rest of the paper is structured as follows. Section 2 introduces the model structure representation proposed for this work. Section 3 gives details about the GP program encoding as well as the definition of the multiobjective fitness function. Section 4 presents the simulation and results of the chaotic system analysis.

0-7803-5536-9/99/$10.00 ©1999 IEEE

Finally, Section 5 draws some conclusions and future developments of this alternative for chaotic modelling.

2 Difference Equation Program Model Representation

The model representation used is the NARMAX model (Leontaritis and Billings, 1985). Previous work has formulated a non-linear system identification tool based upon the NARMAX model that uses GP for determining the appropriate non-linear model structure, which has been successfully applied to simulated and real identification problems (Rodríguez-Vázquez, *et al.*, 1997; Rodríguez-Vázquez and Fleming, 1998). The general NARMAX model is defined as a non-linear function of the input, output and noise signal terms. By selecting the input terms in the NARMAX model formulation to be zero gives the Non-linear Auto Regressive Moving Average (NARMA) model (Leontaritis and Billing, 1985) description

$$y(k) = F^{\ell}\{y(k-1),\cdots,y(k-n_y),e(k-1),\cdots,e(k-n_e)\}+e(k) \quad (1)$$

where n_y and n_e are the maximum lags considered for the process and noise terms, respectively. Moreover, $\{y(k)\}$ denotes a measured time series or signal, $\{e(k)\}$ is an unobservable zero mean independent sequence and $F^{\ell}\{\bullet\}$ is some non-linear function. As has been stated in previous work (Chen and Billings, 1989), the most typical choice for $F^{\ell}\{\bullet\}$ in equation (1) is a polynomial expansion. The model is linear in the parameters and can, therefore, be estimated by means of a Least-Squares algorithm.

This model representation is used to model the dynamic behaviour of chaotic systems. The next section describes the integration of this representation into the GP algorithm and describes the formulation of the multiobjective fitness function.

3 Multiobjective Genetic Programming

3.1 Genetic Programming Encoding

Based upon the model representation defined by equation (1), the GP population consists of tree-structured individuals that readily represent alternative structures for the application of the NARMA approach. Potential models are encoded as hierarchical tree structures, thus providing a dynamic and variable representation, and these constitute members of a population of different model structures. These structures consist of functions (internal nodes) and terminals (leaf nodes) appropriate to the problem domain. Hence, the function set is here defined as F = {ADD, MULT} = {+, *}, and the terminal set as T = $\{X_o, ..., X_{ny}, X_{ny+1}, ..., X_{ny+ne}\}$ = {c, y(k-1), ..., y(k-n_y), e(k-1), ..., e(k-n_e)}. An example of this hierarchical tree representation of the polynomial NARMA model is expressed in Polish notation as (ADD (ADD X1 X4) (MULT (ADD X2 X3)(ADD X1 X2))). This is equivalent to the polynomial non-linear model defined as

$$y(k)=\theta_0+\theta_1 y(k-1)+\theta_2 y(k-2)+\theta_3 e(k-1)+ \theta_4 y(k-1)^2+\theta_5 y(k-1)y(k-2) \quad (2)$$

where {X0, X1, X2, X3} = {1.0 (the constant term), y(k-1), y(k-2), e(k-1)}. A Least-Squares algorithm is applied to compute the parameter vector θ_i to minimise the residual of errors ε between the measured output $y(k)$ and the predicted output $\hat{y}(k)$ that is given by

$$\varepsilon(k)=y(k)-\hat{y}(k,\hat{\theta}) \quad (3)$$

This parameter estimation algorithm works by first calculating the process terms coefficients and, using equation (3), the residuals are computed. Once the residuals are known, these are incorporated into the model and a new set of parameters is estimated.

3.2 Multiobjective Fitness Function

In order to perform selection, all model measures considered in the identification are evaluated for each member of the population. The fitness value of each population member is assigned by means of a rank-based fitness method (Fonseca and Fleming, 1993). This fitness evaluation is based on the definition of Pareto-optimality or nondominance. If we consider a minimisation problem and, given two n components objective function vectors, $\vec{f}_u$ and $\vec{f}_v$, we can say that $\vec{f}_u$ dominates $\vec{f}_v$ (is Pareto-optimal) if

$$\forall i \in \{1,\ldots,n\}, f_{u_i} \le f_{v_i} \wedge \exists i \in \{1,\ldots n\}, f_{u_i} < f_{v_i} \quad (4)$$

producing a set of possible and valid solutions known as the Pareto-optimal or nondominated set. Selection in the evolutionary process is made using a method of ranking which favours non-dominant members of the population (Fonseca and Fleming, 1993). Thus, the model complexity, performance and validation attributes of the modelling process can simultaneously be evaluated by considering them in the multiobjective fitness function. The objectives involved are defined and classified as shown in Table 1.

Regarding the validation criteria, Billings and Tao (1991) have introduced time series validation tests based upon general correlations. These are given by

$$\begin{cases} \Phi_{\varepsilon'\varepsilon'}(\tau) = E\left[\left(\varepsilon(t) - \overline{\varepsilon(t)}\right)\left(\varepsilon(t-\tau) - \overline{\varepsilon(t)}\right)\right] = \delta(\tau) \\ \Phi_{\varepsilon'(\varepsilon'^2)}(\tau) = E\left[\left(\varepsilon(t) - \overline{\varepsilon(t)}\right)\left(\varepsilon^2(t-\tau) - \overline{\varepsilon^2(t)}\right)\right] = 0 \\ \Phi_{(\varepsilon'^2)(\varepsilon'^2)}(\tau) = E\left[\left(\varepsilon^2(t) - \overline{\varepsilon^2 t}\right)\left(\varepsilon^2(t-\tau) - \overline{\varepsilon^2(t)}\right)\right] = \delta(\tau) \end{cases} \tag{5}$$

The objectives related to the performance of candidate NARMA models were defined shown in Table 1. However, due to the unpredictable feature of chaos, long-term prediction was not selected as an objective, and was ignored.

The correlation objective functions were cast as constraints. The target value to be attained was given by the 95% confidence limit. This restriction was assumed in order to identify valid NARMA models. The correlation-based validation objectives were ($2*\tau + 1$) elements vector. In order to define these functions as scalars, the following operation was assumed

$$CCF = \max\left|abs\left(\Phi_{ab}(\tau)\right)\right|. \tag{6}$$

where a,b = ε or $(\varepsilon)^2$. In the case of the autocorrelation of the residuals and the second higher-order correlation function, $\Phi ab(\tau) = 0$, for $\tau = 0$.

Table 1. Description of the objectives considered in the MOGP-identification procedure.

Attribute	Objective	Description
Model complexity	Model size	Number of process and noise terms
	Model degree	Maximum order term
	Model Lag	Maximum lagged input, output and noise terms
Model Performance	Residual variance	Variance of the predictive error between the OSAPE
	Long-term prediction error	Variance of the LTPE
Model validation	Equation (5)	Correlation based test functions for model with noise additive at the output

4 Simulation Results

This section discusses the identification of non-linear polynomial models from chaotic data sets. The chaotic system used to test the applicability of the NARMA-GP approach was Chua's circuit (Chua *et al.*, 1986). This system is one of the most popular benchmarks for studying non-linear oscillations.

The main reasons for Chua's circuit being the most well studied non-linear circuit are i) it is a simple and quite robust circuit that can become chaotic, ii) the chaotic behaviour of this circuit has been observed by computer simulation, iii) mathematical studies have confirmed the circuit chaotic nature and iv) it exhibits a variety of non-linear dynamics and can therefore be considered as a prototype model of chaos.

The normalised equations of Chua's circuit (Figure 1) can be written as (Chua *et al.*, 1986)

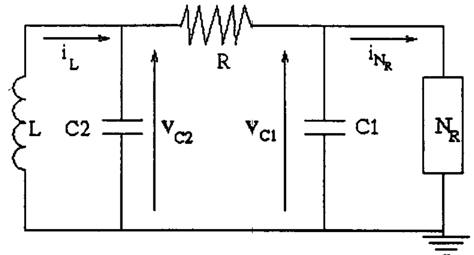


Figure 1. Chua's circuit.

$$\begin{cases} \dot{x} = \alpha(y - h(x)) \\ \dot{y} = x - y + z \\ \dot{z} = -\beta y \end{cases} \qquad h(x) = \begin{cases} m_1 x + (m_0 + m_1) & x \geq 1 \\ m_0 x & |x| \leq 1 \\ m_1 x - (m_0 + m_1) & x \leq -1 \end{cases} \tag{7}$$

where $m_0 = -1/7$ and $m_1 = 2/7$. The variation of α and β parameters drives the system to display several regular and chaotic regimes. The well-known double-scroll attractor is obtained for α=9.0 and β=100/7. Based upon these values, equation (7) was simulated using a Runge-Kutta of 4th-order with step size of 0.001[1]. The z-coordinate of the double-scroll attractor was then chosen for this example (see Figure 2). The identification technique described in Rodríguez-Vázquez and Fleming (1998) and defined for time series prediction in this work, was used to identify non-linear polynomial models of the form of equation (1).

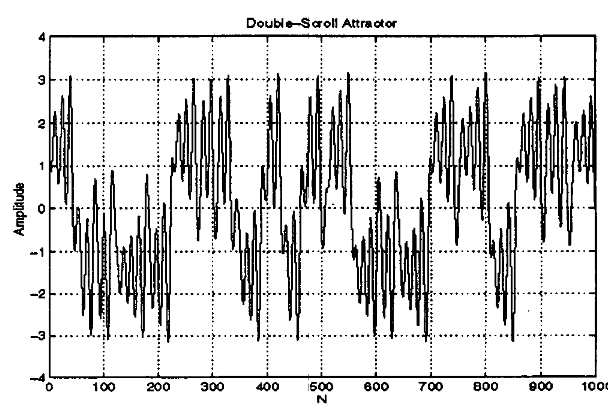


Figure 2. Z-coordinate of Chua's circuit that exhibits chaotic behaviour.

[1] We wish to acknowledge the work of Mendes (1995) who performed this simulation and provided the data.

Because some information about the dynamical non-linearity of the system is known *a priori*, the number of fixed points (equilibrium) is introduced as an objective since it has a direct relation to the dynamic model degree.

The concept of fixed point can be a way to determine the structure of the non-linear polynomial model, as pointed out by Mendes (1995). The term clusters and cluster coefficients (see Definition 2 below) are useful tools that are used to obtain not only the number of fixed points of a system but also the location of such points.

4.1 Term Clustering

The NARX model, the deterministic part of the aforementioned NARMAX model, can be expanded as the summation of terms with degrees of nonlinearity in the range $1 \le m \le l$. Each m th-order term can contain a p th-order factor in $y(k-n_i)$ and a $(m-p)$ th-order factor in $u(k-n_i)$ and multiplied by a coefficient $C_{p,m-p}(n_1, \ldots, n_m)$ as follows

$$y(k) = \sum_{m=0}^{l} \sum_{p=0}^{m} \sum_{n_1,n_m}^{n_y,n_u} C_{p,m-p}(n_1,\ldots,n_m) \prod_{i=1}^{p} y(k-n_i) \prod_{i=p+1}^{m} u(k-n_i) \quad (8)$$

where,

$$\sum_{n_1,n_m}^{n_y,n_u} \equiv \sum_{n_1=1}^{n_y} \cdots \sum_{n_m=1}^{n_u} \quad (9)$$

and the upper limit is n_y if the summation refers to factors in $y(k-n_i)$ or n_u for factors in $u(k-n_i)$.

Definition 2. Cluster Coefficients (Mendes, 1995).

The constants $\sum_{n_1,n_m}^{n_y,n_u} C_{p,m-p}(n_1,\ldots,n_m)$ are the coefficients of the *term clusters* $\Omega_{y^p u^{m-p}}$, which contain terms of the form $y(k-i)^p u(k-j)^{m-p}$ for $m=0, \ldots, l$ and $p=0, \ldots, m$. Such coefficients are called cluster coefficients represented as $\sum_{y^p u^{m-p}}$.

From the last definition, one can say that the set of all candidate terms for a NARX model is the union of all possible clusters up to degree l. That is,

$$\{\text{All possible terms}\} = \bigcup_{\substack{p=0,\ldots,m \\ m=0,\ldots,l}} \Omega_{y^p u^{m-p}}$$

$$= \text{constants} \cup \Omega_y \cup \Omega_u \cup \Omega_{y^2} \cup \Omega_{yu} \cup \Omega_{u^2} \cup \ldots$$

$$\ldots \cup \text{all possible combinations up to degree } l.$$

4.2 Fixed Points

Generally, a fixed point of a system is defined as the point where $y(k)=y(k+i)$, $i \in \mathbb{Z}$. The fixed points will be calculated for the autonomous polynomial of the system under study. If the original polynomial is non-autonomous, then $u(k-i)$, $i=0,1,\ldots$ is set to be zero so that the only remaining terms involve the output. Thus, the possible clusters of an autonomous polynomial with degree of nonlinearity l are Ω_0=constant, Ω_y, Ω_{y^2}, . . ., Ω_{y^l}, and the fixed points are the roots of the *clustered polynomial* expressed as,

$$y(k) = C_{0,0} + y(k) \sum_{n_1=1}^{n_y} C_{1,0}(n_1) + y(k)^2 \sum_{n_1,n_2}^{n_y,n_y} C_{2,0}(n_1,n_2)$$

$$+ \ldots + y(k)^l \sum_{n_1,\ldots,n_l}^{n_y,\ldots,n_y} C_{l,0}(n_1,\ldots,n_l) \quad (10)$$

From the definition of *cluster coefficients*, equation (10) can be rewritten as,

$$\Sigma_{y^l} y^l(k) + \ldots + \Sigma_{y^2} y^2(k) + \Sigma_y y(k) + \Sigma_0 = 0 \quad (11)$$

where $\Sigma_0 = C_{0,0}$ is a constant. Equation (11) shows that the degree of nonlinearity of the autonomous polynomial gives the numbers of fixed points if $\Sigma_y^1 \neq 0$.

Although *a priori* information is available about the system structure, a disadvantage that the MOGP framework exhibits is the inability to deal with equality constraints. Therefore, in order to overcome this weakness, the objective related to the model degree was redefined as

$$Obj_{DEG} = abs|DEG - No\_FXP| \quad (12)$$

where *No_FXP*, the number of fixed points, is defined to be 3 as shown in Fig. 2.

4.3 Simulation

The multiobjective genetic programming method was run with a population of 100 individuals for 100 generations. Crossover and mutation probabilities were 90% and 10 %. The function set was as defined in the aforementioned example of hierarchical tree encoding of NARMA models, and the terminal set consisted of the last 10 values of the signal. Thus,

$$\mathcal{T} = \{y(k-1), \ldots, y(k-10)\}$$

A non-dominated set of chaotic models obtained from one out of several MOGP runs is shown in Table 2 It is relevant to mention that MOGP can produce, at each run, a set of equivalent models with 'good' model attributes.

Table 2. Details of the set of dynamical models used to reconstruct the double-scroll attractor plotted in Figure 4.

Model	p	DEG	LAG	OSAPEx10^{-5}	Fixed Point Location
1	9	3	3	3.5222	(-1.5323, 0, 1.5323)
2	10	3	4	3.1768	(-1.5341, 0, 1.5341)
3	12	3	4	1.3427	(-1.5096, 0, 1.5096)
4	13	3	4	1.2139	(-1.5013, 0, 1.5013)
5	14	3	4	1.2133	(-1.5023, 0, 1.5023)
6	14	3	5	1.1768	(-1.5028, 0, 1.5028)
7	14	3	8	1.1668	(-1.5048, 0, 1.5048)

It is desirable for the validation of chaotic models that the fixed points of the estimated models should be as close as possible to the fixed points of the original system. Based on the concept of symmetry of fixed points, Mendes (1995) stated that the fixed points of a cubic polynomial with a (-z, 0, z) symmetry are obtained from the cluster polynomial (see equation 11) where $\Sigma_0 = 0$. This fact guarantees that the respective dynamical polynomial model should not have any terms taken from the clusters Ω_0 and Ω_{y^2}. In this case the symmetrical fixed points are at

$$\bar{y} = \pm\sqrt{\frac{\Sigma_y - 1}{\Sigma_{y^3}}} \qquad (13)$$

It is interesting to note that the double-scroll attractor possesses these characteristics. It is therefore seen from the results that the dynamical polynomial models identified by using GP only possess linear and cubic terms in their structures. Table 2 also provides details of the location of the fixed points for each model.

An additional feature for the validation of chaotic models is to verify if such models settle to attractors which resemble the geometry of the original data. Thus, the embedded attractors of the identified models are shown in Fig. 4.

From these results, the most parsimonious model that was identified and can reproduce the double-scroll attractor consists only of nine terms. The structure is then given as

$$\begin{aligned} z(k) = {} & 2.9579z(k-1) - 2.9369z(k-2) + 1.0354z(k-3) \\ & -0.6932z(k-1)^3 + 2.3307z(k-1)^2 z(k-2) \\ & -1.1913z(k-1)^2 z(k-3) - 1.9134z(k-1)z(k-2)^2 \\ & +1.9162z(k-1)z(k-2)z(k-3) \\ & -0.4732z(k-1)z(k-3)^2 \end{aligned} \qquad (14)$$

The validation correlation tests described in equation (5) applied to this model are illustrated in Fig. 3.

5 Conclusions and Future Perspectives

Genetic programming is a technique which has been developed for the purpose of solving certain classes of optimisation problem Combining this technique with the general NARMAX representation has provided a powerful tool for non-linear system identification.

The applicability of this technique has been tested, in this work, on a non-linear system exhibiting chaotic motion.

Under chaotic conditions, the well-known benchmark problem, Chua's circuit, was studied. Chaos is easy to postulate but hard to diagnose and to identify. As suggested by Casdagli *et al.* (1992), GP may be useful in forecasting chaotic series and offers an alternative for this kind of identification problem. Hence, GP was used here to identify models that reproduce the chaotic behaviour presented in these two aforementioned systems.

Considering validation and additional information extracted from the original data prior to the identification process, MOGP was able to manage the search space and extract models that could reproduce the dynamics presented in the system under investigation.

One feature of chaos is its long-term unpredictability. For this reason, the long-term prediction error criterion was "ignored" in the multiobjective function definition. For the case of Chua's circuit, the set of models bred using MOGP was able to reproduce attractors which resemble the geometry of the original data.

Further work could include not only the evaluation of statistical but also dynamical validation tools such as the Lyapunov exponent and correlation dimension (Mendes, 1995). These criteria could add a degree of selectivity to the identification procedure and point to more accurate models.

In view of the results of the paper, evolutionary computing methods, in general, provide an alternative for generating simple and near-optimal solutions of practical and complex problems.

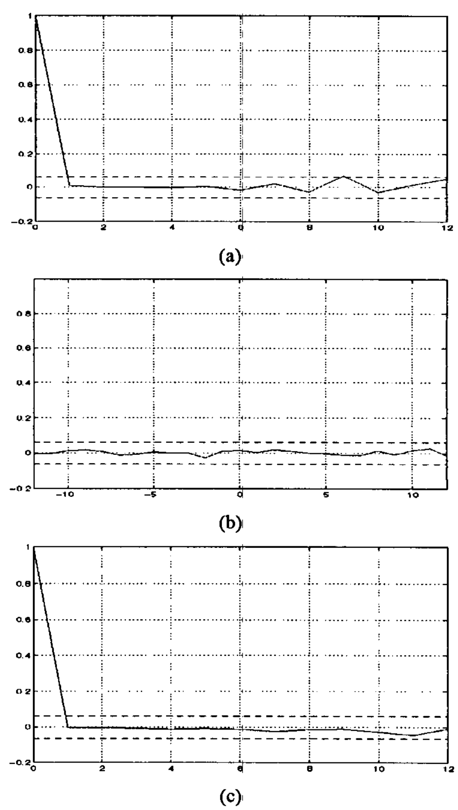
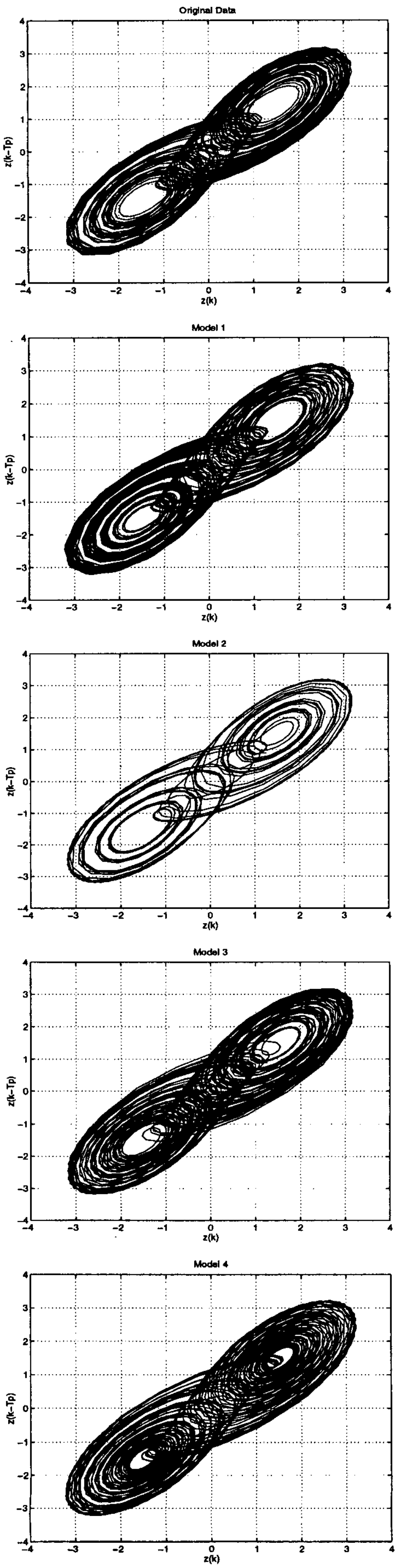


Figure 3. Correlation tests performed in the Chua's circuit data generated from model 1. (a) $\Phi_{\varepsilon'\varepsilon'}(\tau)$, (b) $\Phi_{\varepsilon'\varepsilon^{2'}}(\tau)$, (c) $\Phi_{\varepsilon^{2'}\varepsilon^{2'}}(\tau)$.



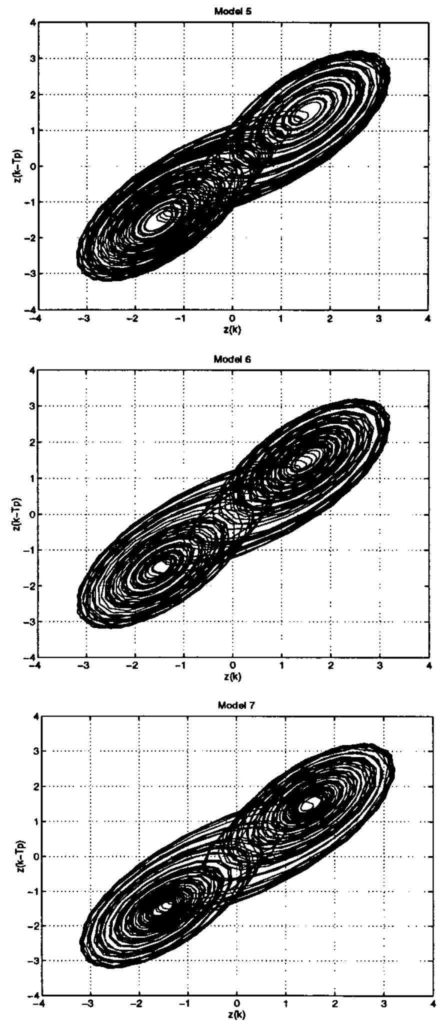


Figure 4. Double-scroll attractor reconstructed using models identified by means of the NARMA-GP approach.

ACKNOWLEDGEMENTS

Katya Rodríguez-Vázquez gratefully acknowledges the financial support of Consejo Nacional de Ciencia y Tecnología (CONACyT), México.

BIBLIOGRAPHY

BILLINGS, S.A. AND Q.H. TAO (1991) Model Validity Tests for Non-Linear Signal Processing Applications. *Int. J. Control*, **54**(1), pp. 157-194.

CASDAGLI, M., D. DES JARDINS, S. EUBANK, J.D. FARMER, J. GIBSON AND J. THEILER. Non-Linear Modeling of Chaotic Time Series: Theory and Applications. In *Applied Chaos* (J.H. Kim and J.S. Stringer, Eds.). Wile Interscience.

CHEN, S. AND S.A. BILLINGS (1989) Representation of Non-Linear Systems: the NARMAX Model. *Int. J. Control*, 49(3), pp. 1013-1032.

CHUA, L.O., M. KAMURO AND T. MATSUMOTO (1986) The Double Scroll Family. *IEEE Trans. Circuits Syst.*, **33**(11), pp. 1072-1118.

FONSECA, C.M. AND P.J. FLEMING. (1993) Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization. *Proc. 5th Int. Conf. Genetic Algorithms*, pp. 416-423.

IBA, H. T. KURITA, H. DE GARIS AND T. SATO (1993) System Identification Using Structured Genetic Algorithms. In *Proc. 5th Int. Conf. Genetic Algorithms*, pp. 467-491.

IVANKHNENKO, A.G. (1971) Polynomial Theory of Complex Systems, *IEEE Trans. Systems, Man and Cybernetics*, **1**(4), 364-378.

KOZA, J.R. (1992) *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press.

LEONTARITIS, I.J. AND S.A. BILLINGS (1985) Input-Output Parametric Models for Non-Linear Systems. Part I and Part II. *Int. J. Control*, 41(2), pp. 304-344.

MENDES, E.M.A.M. (1995) Identification of Non-Linear Discrete Systems with Intelligent Structure Detection. PhD Thesis. Department of Automatic Control and Systems Engineering. University of Sheffield, U.K.

MULLOY, B.S., R.L. RIOLO AND R.S. SAVIT (1996) Dynamics of Genetic Programming and Chaotic Time Series Prediction. In (Koza *et al.*, 1996). pp. 157-165.

OAKLEY, E. H. N. (1994) Two Scientific Applications of Genetic Programming: Stack Filters and Non-Linear Equation Fitting to Chaotic Data. In (Kinnear, 1994), pp. 369-389.

RODRÍGUEZ-VÁZQUEZ, K., C.M. FONSECA AND P.J. FLEMING. (1997) An Evolutionary Approach to Non-Linear Polynomial System Identification. In *11th IFAC Symposium on System Identification*, Vol. **3**, pp. 1395-1400.

RODRÍGUEZ-VÁZQUEZ, K. AND P.J. FLEMING (1998) Multiobjective Genetic Programming for a Gas Turbine Engine Model Identification. *Proc UKACC Int. Conf. on CONTROL '98*, pp. 1385-1390.