Smart Structures and Systems, *Vol. 15, No. 3 (2015) 897-911*
DOI: http://dx.doi.org/10.12989/sss.2015.15.3.897

Evolutionary computational approaches for data-driven modeling of multi-dimensional memory-dependent systems

Ali Bolourchi* and Sami F. Masri[a]

Viterbi School of Engineering, University of Southern California, 3620 South Vermont Avenue, KAP 210, Los Angeles, CA, 90089-2531 USA

(Received May 30, 2014, Revised August 20, 2014, Accepted August 30, 2014)

Abstract. This study presents a novel approach based on advancements in Evolutionary Computation for data-driven modeling of complex multi-dimensional memory-dependent systems. The investigated example is a benchmark coupled three-dimensional system that incorporates 6 Bouc-Wen elements, and is subjected to external excitations at three points. The proposed technique of this research adapts Genetic Programming for discovering the optimum structure of the differential equation of an auxiliary variable associated with every specific degree-of-freedom of this system that integrates the imposed effect of vibrations at all other degrees-of-freedom. After the termination of the first phase of the optimization process, a system of differential equations is formed that represent the multi-dimensional hysteretic system. Then, the parameters of this system of differential equations are optimized in the second phase using Genetic Algorithms to yield accurate response estimates globally, because the separately obtained differential equations are coupled essentially, and their true performance can be assessed only when the entire system of coupled differential equations is solved. The resultant model after the second phase of optimization is a low-order low-complexity surrogate computational model that represents the investigated three-dimensional memory-dependent system. Hence, this research presents a promising data-driven modeling technique for obtaining optimized representative models for multi-dimensional hysteretic systems that yield reasonably accurate results, and can be generalized to many problems, in various fields, ranging from engineering to economics as well as biology.

Keywords: computational intelligence; genetic algorithms; differential equations; hysteretic behavior; data-driven modeling; identification; multi-dimensional systems; genetic programming

1. Introduction

Modeling and analysis of complex multi-dimensional nonlinear memory-dependent systems is a broad research area with applications in many fields such as mechanics, economy, aeronautics, amongst others. Over the past several decades, nonlinear systems incorporating memory-dependent dissipative phenomena have been investigated in many studies, mainly through parametric or nonparametric modeling (Kerschen *et al.* 2006). Parametric modeling is a class of techniques used to characterize hysteretic systems by attributing a suitable model to the

*Corresponding author, Ph.D., E-mail: bolourch@gmail.com
[a] Professor, E-mail: masri@usc.edu

http://www.techno-press.org/?journal=sss&subpage=8 ISSN: 1738-1584 (Print), 1738-1991 (Online)

system, and identifying the model parameters or detecting their changes (Bouc 1967, Wen 1976 Smyth *et al.* 1999, Chatzi *et al.* 2010). However, the success of parametric modeling highly relies on the user providing an accurate representative model based on a clear understanding of the physical phenomena. Non-parametric modeling, on the other hand, carries out the identification automatically, without setting any major hypothesis about the system upfront. In this category, Neural Networks are powerful tools of modeling complex systems involving memory-dependent characteristics (Bani-Hani and Ghaboussi 1998, Pei and Smyth 2006a, b, Zhao and Tan 2008). Their data driven analysis is suitable when no additional information about the underlying systems is available. However, Neural Networks process data in a completely blind manner, that is not suitable for revealing the physics behind complex systems. Therefore, they leave no room for further investigation of the resultant models, and to provide physical insight into the systems under investigation. Polynomial-based approaches are another class of nonparametric techniques that are extensively employed for the modeling of nonlinear phenomena in dynamic environments (Masri and Caughey 1979, Masri *et al.* 2004, 2006a, b, Tasbihgoo *et al.* 2007). However, specifying the order of polynomials and the composition of basis functions requires insight into the dynamics of the problem. Moreover, polynomial approximations of systems exhibiting sharp corners in their response results in undesired Gibbs phenomena.

Consequently, there is currently a lack of intelligent data-driven computational techniques that provide robust physics-based computational models, that represent multi-dimensional dynamical systems with hysteresis effect. Evolutionary computational methodologies are inspired by natural phenomena to provide elegant solutions to complex real-world problems in various fields. In this class of problem-solving techniques, Genetic Programming (GP), is built on evolutionary algorithms, and offers a great potential for the identification of complex dynamical systems exhibiting non-conservative dissipative behavior. GP has proven to be successfully applicable to various classes of problems in different fields (Tackett 1993, Gruau 1994, Howard and Roberts 2002, Becker *et al.* 2007, Schmidt and Lipson 2009, Alavi *et al.* 2010, Silva *et al.* 2011, Gandomi and Alavi 2011). To our knowledge, GP has never been adapted to be employed for the modeling of multi-dimensional systems with hysteretic behavior. Therefore, this investigation explores the potential of evolutionary approaches to discover the system of differential equations that govern the behavior of complex nonlinear multi-dimensional hysteretic systems.

A general procedure concerning the modeling of complex one-dimensional systems associated with challenging type of nonlinearities was presented in Bolourchi (2014) and Bolourchi *et al.* (2015). That procedure is extended in this paper and is implemented for the modeling and analysis of non-linear multi-dimensional systems with memory-dependent dissipative characteristics. The eventual aim is to utilize GP as the main problem-solving engine for discovering the suitable "*structure*" of the differential equations that govern the behavior of multi-dimensional hysteretic systems. GP is also coupled with stochastic parameter optimization techniques, employing Genetic Algorithms, to optimize the "*parameters*" embedded in the differential equations based on the global performance of the integrated GP-found structures in the system of differential equations for estimating the response of the underlying multi-dimensional hysteretic system.

It is shown that the suggested identification methodology provides reduced-complexity systems of differential equations that govern the dynamics of complex nonlinear multi-dimensional systems with non-conservative dissipative traits. Several test cases are provided to assess the applicability, reliability, and validity of the methodology.

This paper is organized as follows: the formulation of the investigated system, the corresponding assumptions, and the details of the excitations and responses are provided in

Section 2; an introduction to GP within the scope of the investigation in this paper is presented in Section 3; the proposed evolutionary-based approach for the modeling of multi-dimensional hysteretic systems is presented in Section 4; the results of the introduced modeling procedure is presented in Section 5; in-depth interpretation of the findings are discussed in Section 6; advantages and challenges of the approach are described in Section 7; and the conclusion is provided in Section 8.

2. Multi-dimensional hysteretic systems

2.1 Investigated models

The identification method under discussion is implemented on a benchmark multi-degree-of-freedom (MDOF) system exhibiting non-conservative nonlinear behavior. This structure has been employed for testing the robustness of other identification schemes in the past (Masri *et al.* 1987, Smyth *et al.* 2002, Masri *et al.* 2005). This three degree-of-freedom system in Fig. 1 is composed of three unequal lumped masses that are linked by six arbitrary components to all other masses and to a fixed support.

In Fig. (1), x_i is the displacement, $f_i(t)$ is external excitation, and m_i is the mass at DOF i. The semi-physical Bouc-Wen model is well studied in the literature to represent non-conservative dissipative systems (Smyth *et al.* 2002). The interconnecting elements in the structure above, denoted as g_i, are governed by the Bouc-Wen model, and consequently, account for the hysteretic behavior of the MDOF system. The Bouc-Wen model for a single degree-of-freedom system is formulated as follows

$$f(t) = m\ddot{x} + r(x,\dot{x}) \tag{1}$$

$$r(x,\dot{x}) = \frac{1}{\eta}\left[A\dot{x} - \upsilon\left(\beta|\dot{x}||r(x,\dot{x})|^{n-1}r(x,\dot{x}) - \gamma\dot{x}|r(x,\dot{x})|^{n}\right)\right] \tag{2}$$

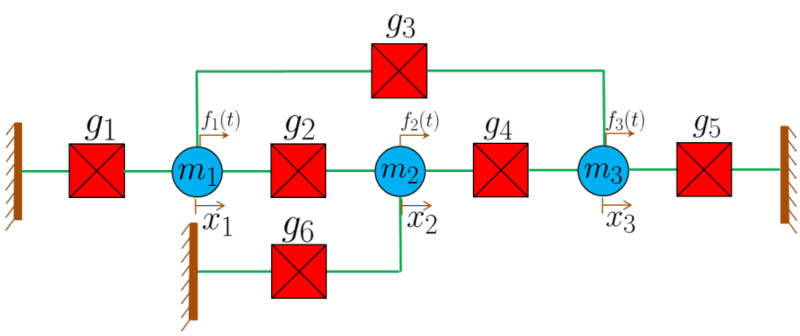

Fig. 1 The benchmark 3 degree-of-freedom system with hysteretic interconnecting links g_i s

where $f(t)$ is the external excitation, m is the lumped mass, $\ddot{x}$ is the acceleration, $r(x,\dot{x})$ is the restoring force, x is the displacement, and $\dot{x}$ is the velocity. The shape of the hysteretic loops is ruled by the parameters of the model: $\eta, \mathrm{A}, \nu, \beta, \gamma$ and n (Smyth *et al.* 2002). Therefore, the MDOF system of Fig. 1 under discussion is formulated as follows

$$\begin{cases} f_1(t) = m_1\ddot{x}_1 + g_1(u_1,\dot{u}_1) + g_2(u_2,\dot{u}_2) + g_3(u_3,\dot{u}_3) & 3(\mathrm{a}) \\ f_2(t) = m_2\ddot{x}_2 + g_2(u_2,\dot{u}_2) + g_4(u_4,\dot{u}_4) + g_6(u_6,\dot{u}_6) & 3(\mathrm{b}) \\ f_3(t) = m_3\ddot{x}_3 + g_3(u_3,\dot{u}_3) + g_4(u_4,\dot{u}_4) + g_5(u_5,\dot{u}_5) & 3(\mathrm{c}) \\ \dot{g}_1(u_1,\dot{u}_1) = \dot{u}_1 - |\dot{u}_1| g_1(u_1,\dot{u}_1) + 0.5u_1 |g_1(u_1,\dot{u}_1)| & 3(\mathrm{d}) \\ \dot{g}_2(u_2,\dot{u}_2) = \dot{u}_2 - |\dot{u}_2| g_2(u_2,\dot{u}_2) + 0.5u_2 |g_2(u_2,\dot{u}_2)| & 3(\mathrm{e}) \\ \dot{g}_3(u_3,\dot{u}_3) = \dot{u}_3 - |\dot{u}_3| g_3(u_3,\dot{u}_3) + 0.5u_3 |g_3(u_3,\dot{u}_3)| & 3(\mathrm{f}) \\ \dot{g}_4(u_4,\dot{u}_4) = \dot{u}_4 - |\dot{u}_4| g_4(u_4,\dot{u}_4) + 0.5u_4 |g_4(u_4,\dot{u}_4)| & 3(\mathrm{g}) \\ \dot{g}_5(u_5,\dot{u}_5) = \dot{u}_5 - |\dot{u}_5| g_5(u_5,\dot{u}_5) + 0.5u_5 |g_5(u_5,\dot{u}_5)| & 3(\mathrm{h}) \\ \dot{g}_6(u_6,\dot{u}_6) = \dot{u}_6 - |\dot{u}_6| g_6(u_6,\dot{u}_6) + 0.5u_6 |g_6(u_6,\dot{u}_6)| & 3(\mathrm{i}) \end{cases}$$

where u_i and $\dot{u}_i$ are the relative displacement and the relative velocity of the two ends of the interconnecting link g_i. $\dot{u}_i$ is the derivative of u_i which is determined according to the topology of the system: $u_1 = x_1$, $u_2 = x_2 - x_1$, $u_3 = x_3 - x_1$, $u_4 = x_3 - x_2$, $u_5 = x_5$, $u_6 = x_6$.

For convenience, the parameters of the Bouc-Wen elements in all interconnecting components are identical, and presented in Table 1. Note that this system is more complex than regular MDOF systems, typically encountered in civil structures, in which characteristic matrices, such as stiffness and damping, are simplified because every mass is only connected to its adjacent masses. Thus, the non-parametric identification of this complex system to yield surrogate representative models is a challenging problem, and an ideal example to evaluate the capabilities of new modeling techniques. Lumped masses of the MDOF system under investigation are stimulated by dissimilar zero-mean stationary Gaussian white noise excitations. The applied training excitations at each degree-of-freedom (DOF) i have identical statistical properties, which are presented in Table 1, and are produced using different random numbers. The external excitations stimulate the system sufficiently strong to disclose the yielding region of the vibrating interconnecting components of the studied system. It is assumed that the external excitations, the lumped masses and their accelerations are known from measurements. Hence, the displacements and velocities can be obtained by careful integration of accelerations. As a result, all states x_i, $\dot{x}_i$ and $\ddot{x}_i$, $i = 1,2,3$, are assumed to be available from measurements.

The proposed non-parametric approach of this research aims to discover an equivalent system of differential equations that describes the three DOF system defined by Eq. (3). This approach benefits from an auxiliary variable that incorporates the effect of all restoring forces from other degrees-of-freedom on a certain mass m_i. Herein, this variable is named the compound restoring force, denoted as r_i at DOF i. Bolourchi (2014) and Bolourchi *et al.* (2015) showed that the presence of the derivative of the restoring force $\dot{r}_i$ in the system of differential equations of SDOF hysteretic systems plays a vital role in providing high-fidelity models for bilinear hysteretic systems. Thus, using the approach of this research, the differential equation associated with each compound restoring force r_i is discovered and optimized. Thus, the target equivalent system of differential equation that is aimed to be discovered is as follows

$$
\begin{cases}
f_1(t) = m_1\ddot{x}_1 + r_1 & 4(a) \\
\dot{r}_1 = q_1(r_1, r_2, r_3, \dot{x}_1, \dot{x}_2, \dot{x}_3, x_1, x_2, x_3) & 4(b) \\
f_2(t) = m_2\ddot{x}_2 + r_2 & 4(c) \\
\dot{r}_2 = q_2(r_1, r_2, r_3, \dot{x}_1, \dot{x}_2, \dot{x}_3, x_1, x_2, x_3) & 4(d) \\
f_3(t) = m_3\ddot{x}_3 + r_3 & 4(e) \\
\dot{r}_3 = q_3(r_1, r_2, r_3, \dot{x}_1, \dot{x}_2, \dot{x}_3, x_1, x_2, x_3) & 4(f)
\end{cases}
$$

where x_i, $\dot{x}_i$ and $\ddot{x}_i$ are the displacement, velocity, and acceleration, respectively, $f_i(t)$ is external excitation, m_i is the mass, and r_i and $\dot{r}_i$ are the compound restoring force and its derivative at DOF i. While Eqs. (4(a)), (4(c)) and (4(e)) are known a priori, the discovered differential equations of the compound restoring forces $\dot{r}_i$ will be placed in Eqs. (4(b)), (4(d)) and (4(e)) to form the surrogate computational model that represent the system under discussion.

3. Introduction to genetic programming

3.1 Overview

Evolutionary Algorithms (EAs) are heuristic optimization methods that mimic the necessary processes for evolution in nature — selection, mutation and crossover — to evolve a population of candidate solutions. Genetic Programming, introduced and popularized by Koza (1992), is a branch of EAs that is capable of evolving any type of structures in the form of expression-trees with quantifiable performance in the domain. (Bolourchi 2014)

Table 1 Properties of the Bouc-Wen model and the applied excitations $f_i(t)$ for training and validation

Parameters of Excitations $f_i(t)$	Values	Bouc-Wen Links g_i	Values
Mean μ	0	n	1.00
Training Standard Deviation σ_t	1.00	A	1.00
Validation Standard Deviation σ_v	1.00	β	1.00
Sampling Frequency $\Delta t/T$	0.05	ν	1.00
Duration	80 Seconds	η	1.00
Masses : m_1, m_2, m_3	0.80, 2.00, 1.20	γ	-0.50

3.2 Population

In this study, the population of evolving candidate solutions are the differential equations of the compound restoring force at each DOF, that will eventually take part in the system of the differential equations of Eq. (4). In order to obtain distinct differential equations, a new population is formed for analyzing each degree-of-freedom. No additional information is assumed to be known about the topology of the system. Thus, the training data set is composed of all displacement and velocity states, $x_1, x_2, x_3, \dot{x}_1, \dot{x}_2, \dot{x}_3$ as well as the compound restoring forces r_1, r_2, r_3 at all degrees-of-freedom. Hence, they form a 9-dimensional function-space domain to estimate the derivative of the restoring force $\dot{r}_i$ at a specific DOF i in Eqs. (4(b)), (4(d)) and (4(e)). Therefore, the MDOF system under investigation has 3 degrees of freedom, but 9 introduced variables constitute 9 dimensions. Note that, due to the complexity of the system, all the available signals are considered as a variable, in the beginning, and consequently, add one dimension to the search domain. However, GP will intelligently remove unrelated variable that do not contribute to the target signal. As a result, dimensionality reduction is conducted automatically throughout the evolution.

In addition to the state variables, algebraic operations ($+, -, \times, /$), as well as the abs and Heaviside $step$ functions are included in the library of essential building blocks, to construct the body of the evolving population. However, only appropriate elements will survive during the course of the evolution to form the most suitable structures, and subsequently, the most accurate differential equations. Then, evolutionary operators (mutation and crossover) are used to advance the evolution. The eventual goal is to discover optimized distinct differential equations for the compound restoring forces of the system to form a complete system of coupled differential equations that fully characterize the system under discussion.

3.3 Fitness criterion

Establishing a suitable fitness criterion is critical for guiding the evolution toward an admissible solution in a timely manner. The fitness error e in GP is calculated using the mean absolute error of the deviation of the model estimate from the target signal. $e = x - \hat{x}$, normalized by the mean absolute value of the reference signal x

$$\varepsilon = \frac{\frac{1}{n}\sum_{i=1}^{n}|\mathrm{x} - \hat{\mathrm{x}}|}{\frac{1}{n}\sum_{i=1}^{n}|\mathrm{x}|} = \frac{\sum_{i=1}^{n}|\mathrm{e}|}{\sum_{i=1}^{n}|\mathrm{x}|} \tag{5}$$

A different fitness criterion is implemented for advancing the parameter optimization by means of Genetic Algorithms. This criterion will be presented later in this paper.

4. Modeling multi-dimensional hysteretic systems: general procedure

The identification procedure consists of two major phases: training and validation.

Table 2 The parameters of Eureqa for obtaining the optimum structures using Genetic Programming

	Eureqa
Parameters	Values
Error Metric (Fitness)	Minimize the Absolute Error
Algebraic Operations	$+ - \times \div$
Basis Functions	*step, abs*
Terminals	Constants, Variables
Stop Criterion	48 hours
Number of Variables	9
Processing Unit	RAM: 16.0 GB; CPU: Intel quad core i7-3370 with hyper threading; CPU clock: 3.40 GHz

Note that due to the incorporation of non-differentiable basis functions, gradient-based methods are not suitable for the optimization process

4.1 Training

Step 1: Excitation

For training, first the synthetic training data set should be generated. Three different, but statistically similar, excitations stimulate the lumped masses at each degree-of-freedom. The properties of the excitations are in accordance with Table 1. They are stationary Gaussian white with zero mean and known standard deviation. The standard deviation controls the intensity of the external excitations. The produced excitations stimulate the lumped masses at all degrees-of-freedom.

Step 2: Response calculation

The reference response of the structure under investigation is calculated by solving the actual system of nonlinear differential equations, described by Eq. (3), using standard time-marching numerical techniques. Then, the compound restoring force at each degree-of-freedom is calculated using the general equation of motion $r_i = f_i(t) - m_i \ddot{x}_i$. The estimate of the derivative of the restoring force is also calculated using two-point finite-difference approximations as follows

$$\dot{r}_i^j = \frac{r_i^{j+1} - r_i^j}{\Delta t} \tag{6}$$

where $\dot{r}_i^j$ is the jth datum in the array of the estimate of the derivative of the restoring force r_i, and r_i^{j+1} and r_i^j are the $j+1$th and jth data in the array of the restoring force r_i, and Δt is the time step.

Step 3: Training data preparation

The derivative of the restoring force $\dot{r}_i$ at DOF i, the restoring force at all degrees-of-freedom r_i, along with displacement and velocity of all degrees-of-freedom form the training data set.

Table 3 Properties of GA for parameter optimization

GA Parameters	Values
Method	**`gaoptimset`**
`PopulationSize`	200
Initial Population	From GP: $\boldsymbol{c}$
`PopInitRange`	$[0.4\mathrm{c} : 2.5\mathrm{c}]$
Bounds (**`lb`**, **`ub`**)	$(0.4\mathrm{c}, 2.5\mathrm{c})$
`TolFun`	0.01
`StallGenLimit`	50

Step 4: Genetic programming computation

Eureqa (Schmidt and Lipson 2009), which is a Genetic Programming toolbox that uses the principles of Genetic Programming to perform this task, was used for the analysis. The fitness is calculated using the absolute error defined earlier by Eq. (5). Algebraic operations and basis functions are the non-terminal building blocks that connect terminal building blocks of trees (ending leafs). Parameters of GP can be found in Table 2.

Step 5: Un-optimized model formation

At the end of three rounds of evolution, three distinct differential equations are obtained for the compound restoring forces at all degrees-of-freedom. Then, they are combined to form the system of differential equations of Eq. (4). However, while the fitness of the obtained models from GP is based on how every single discovered equation can estimate the derivative of the restoring force, the actual performance of the constructed system of coupled differential equations is measured based on their *global* performance. Therefore, since GP is not able to effectively optimize the parameters of the model based on the solution of the differential equation, Genetic Algorithms are employed next to optimize this system of differential equations to improve its global response estimates.

Step 6: Model's parameter optimization

Due to the dependency of the system response on all coupled differential equations combined together, the parameters are optimized using GAs, all together, to enhance the accuracy of the response estimates. Hence, the cost function is defined as the summation of equally-weighted errors between the displacement, velocity and acceleration of the model response and the reference response, at every degree-of-freedom, as follows

$$\varepsilon = \sum_{i=1}^{3} \frac{\|x_i - \hat{x}_i\|}{\|x_i\|} + \frac{\|\dot{x}_i - \hat{\dot{x}}_i\|}{\|\dot{x}_i\|} + \frac{\|\ddot{x}_i - \hat{\ddot{x}}_i\|}{\|\ddot{x}_i\|} \tag{7}$$

where x_i is the displacement, $\dot{x}_i$ is the velocity, and $\ddot{x}_i$ is the acceleration at DOF i and $\hat{x}_i$,

$\hat{\dot{x}}_i$ and $\hat{\ddot{x}}_i$ are their estimates, respectively. GA in Matlab is employed to optimize the parameters, and the assigned options are listed in Table 3. The **gaoptimset** method is used to conduct the evolution of parameters in Matlab. The size of the population is 200.

Step 7: Final optimized model completion

In the end, after the parameter optimization phase is completed, a system of differential equations is obtained whose parameters are also optimized, to accurately represent the MDOF system under investigation, and to yield the best estimates.

4.2 Validation and verification

The last phase of the modeling scheme involves validation and verification by predicting the accuracy of the discovered model when subjected to new excitations whose intensities are substantially different from the training excitations. This phase gages the generalizability of the discovered model, and its applicability in new dynamical environments.

5. Modeling multi-dimensional hysteretic systems: results

This section applies the introduced identification technique that incorporates GP and GA for the identification of a multi-dimensional non-linear hysteretic system with the Bouc-Wen formulation. The model shown in Fig. 1 is excited at all degrees-of-freedom by broad-band uncorrelated forces, described in Table 1, to undergo horizontal motion. The duration of the excitations is 80 seconds, and the time-history of the applied excitation to DOF 2 , as an example, is shown in the lower part of the time-history panels of Fig. 2. The response of the system is obtained by solving the system of coupled differential equations of Eq. (3). The inter-connecting links g_i undergo significant hysteretic deformation. Samples of this behavior for three inter-connecting elements g_i, i=1,2,3 are plotted against the relative displacement of their two ends in Fig. 3.

5.1 Discovering the optimized structure of differential equations by GP

The response of the stimulated 3-DOF system generates 3 batches of training data sets, and are fed to GP to obtain three differential equations associated with the compound restoring force at every degree-of-freedom. It is important to note that none of the information concerning the individual restoring forces g_i s, *i.e.*, neither the formulation of g_i, nor the measurements from a single g_i, is used in the modeling process.

The candidate GP-found differential equations that were obtained after the termination of the optimization for 3 different training datasets are combined to form the entire coupled system of differential equations in Eq. (8). Note that, according to Eq. (3), although 6 Bouc-Wen models are included in the MDOF system of Eq. (3), the estimate of the governing system of differential equations in Eq. (8) involves only three equations for the compound restoring forces. Thus, a fairly simple model is able to represent the investigated complex multi-dimensional system exhibiting nonlinear dissipative memory-dependent behavior.

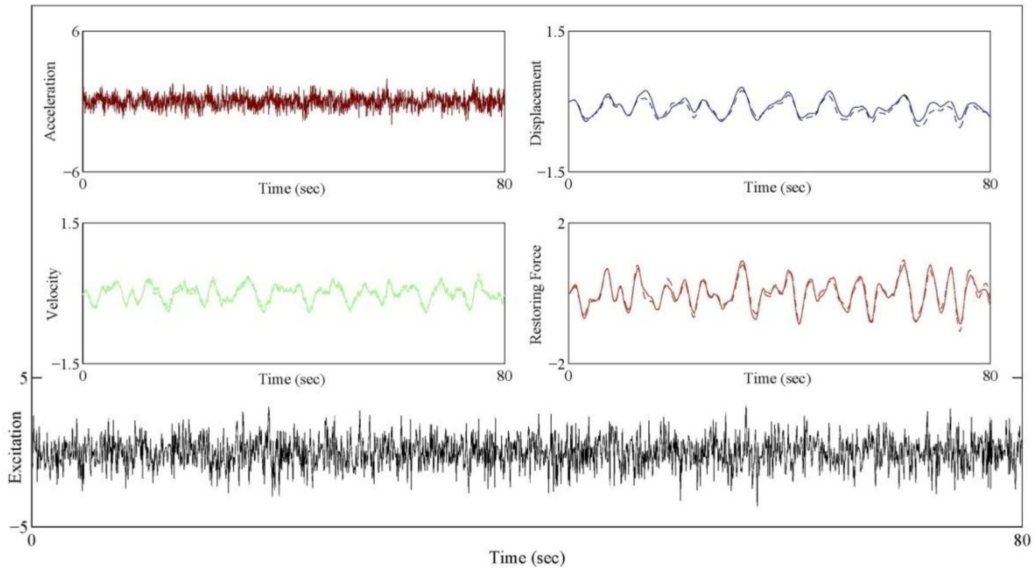

Fig. 2 Comparison of the reference and estimated response of the 3-DOF hysteretic system at DOF 2. The system is stimulated by 3 distinct external training excitations at all degrees-of-freedom. The graphs of the reference and identified response are plotted using solid and dotted lines, respectively

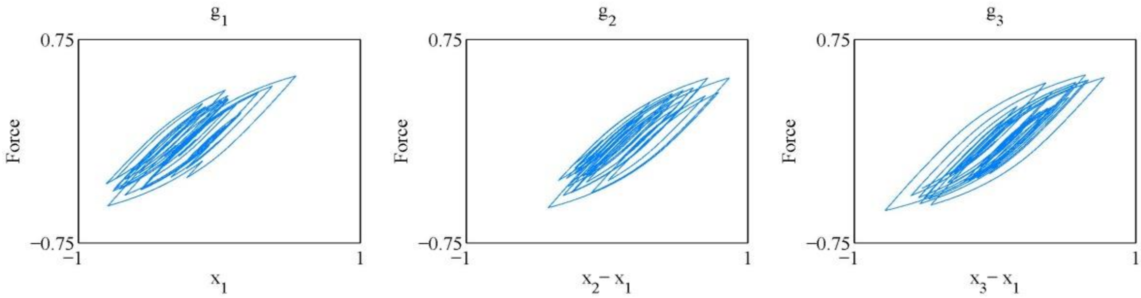

Fig. 3 Phase plots of the force at interconnecting elements of g_1 to g_3 vs. the relative displacement of the corresponding ends, when the 3-DOF hysteretic system is subjected to distinct training excitations with $\sigma_t = 1.0$ at all degrees-of-freedom

$$\begin{cases} \mathrm{f}_1(t) = m_1\ddot{x}_1 + r_1 \\ \dot{r}_1 = 2.797\dot{\mathrm{x}}_1 - \dot{\mathrm{x}}_2 - \dot{\mathrm{x}}_3 - \mathrm{r}_1\mathrm{abs}(\dot{\mathrm{x}}_1) \\ \mathrm{f}_2(t) = m_2\ddot{x}_2 + r_2 \\ \dot{r}_2 = 2.767\dot{\mathrm{x}}_2 - \dot{\mathrm{x}}_1 - \dot{\mathrm{x}}_3 - 0.217\mathrm{r}_2 \\ \mathrm{f}_3(t) = m_3\ddot{x}_3 + r_3 \\ \dot{r}_3 = 2.798\dot{\mathrm{x}}_3 - \dot{\mathrm{x}}_1 - \dot{\mathrm{x}}_2 - \mathrm{r}_3\mathrm{abs}(\dot{\mathrm{x}}_3) \end{cases} \tag{8}$$

5.2 Optimizing the parameters of the system of differential equations by GA

Though the obtained expressions provide the best fit for the defined auxiliary variables $\dot{r}_i$ s, which symbolize the compound restoring forces, the parameters of this system are not optimized by GP to yield accurate response estimates globally. Thus, the weight vector a is introduced to adjust the contribution of every single term in the system of differential equations in such a way that the least difference between the estimate and reference response is achieved. According to Eq. (9), the weight vector has 12 elements, and is optimized by Genetic Algorithms using the error measure defined by Eq. (7). Note that the fact that the weighting parameters are linearly dependent on $\dot{r}_i$ doesn't eliminate the need for a stochastic evolutionary-based optimizer because the cost function of Eq. (7) depends on the response of the *entire* system of a coupled differential equation, rather than the goodness of fit in Eqs. (4(b)), (4(d)) and (4(f)) separately.

$$\begin{cases} f_1(t) = m_1\ddot{x}_1 + r_1 \\ \dot{r}_1 = a(1)2.797\dot{x}_1 - a(2)\dot{x}_2 - a(3)\dot{x}_3 - a(4)r_1 abs(\dot{x}_1) \\ f_2(t) = m_2\ddot{x}_2 + r_2 \\ \dot{r}_2 = a(5)2.767\dot{x}_2 - a(6)\dot{x}_1 - a(7)\dot{x}_3 - a(8)\,0.217r_2 \\ f_3(t) = m_3\ddot{x}_3 + r_3 \\ \dot{r}_3 = a(9)2.798\dot{x}_3 - a(10)\dot{x}_1 - a(11)\dot{x}_2 - a(12)r_3 abs(\dot{x}_3) \end{cases} \quad (9)$$

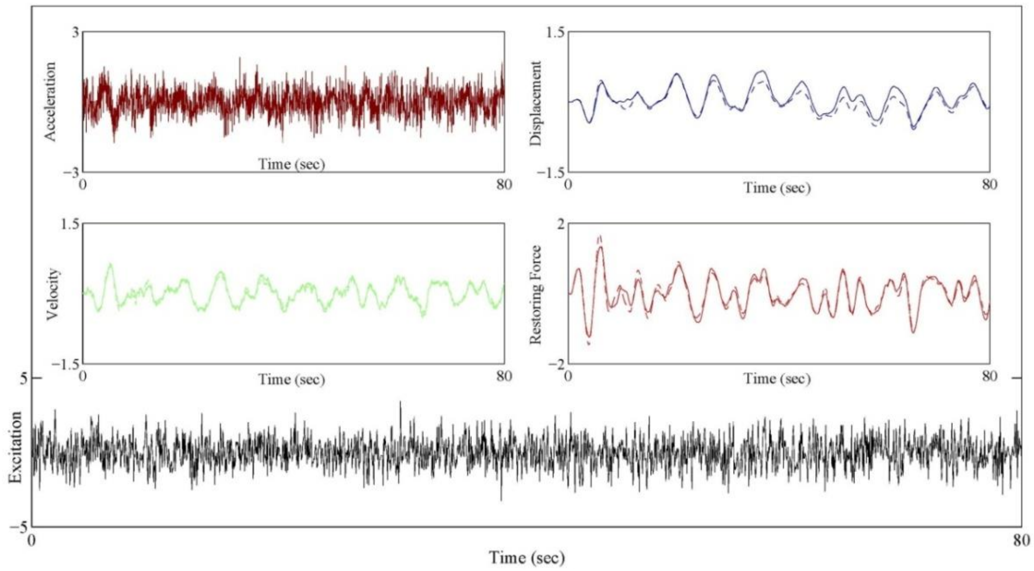

Fig. 4 Comparison of the reference and identified response of the 3-DOF hysteretic system at DOF 2. The system is stimulated by 3 distinct external validation excitations with $\sigma_v = 1.0$ at all degrees-of-freedom which have the same intensity as the training excitation with $\sigma_t = 1.0$. The graphs of the reference and estimated response are plotted using solid and dotted lines, respectively

The cost function associated with the original system of Eq. (9) before optimizing the weights, when $a_i = [111111111111]$, decreases from 266% to 175%, when the optimized weight vector is achieved: $a_f = [0.90, 1.14, 0.90, 0.97, 1.01, 1.12, 0.99, 0.81, 1.07, 0.94, 1.26, 0.90]$. Note that the cost function is defined by Eq. (7).

5.3 Validation of the model

At the end of the training phase, which consisted of both GP for optimizing structures and GA for optimizing parameters, the obtained system of differential equations is validated. The validation is conducted by applying a different random excitation which has the same standard deviation as that of the training, to the MDOF system. Solving the optimized coupled differential equation of Eq. (9) with a_f under the new validation excitation by means of standard numerical techniques provides the response estimate. The time histories of the estimated response and the reference response due to the training excitation are shown in the time history panels of Fig. 4 for DOF 2. Similar results are achieved for other degrees-of-freedom. The reasonable error associated with the estimates confirm that the proposed modeling technique yields reduced-order, reduced-complexity, optimized differential operators that effectively characterize the dynamics of the complex nonlinear MDOF system with hysteretic traits. It is also shown that the model performs well under new dynamical stimulations.

A more comprehensive validation process can be implemented by considering a variety of excitation levels, and conducting a statistical analysis on the outcomes.

6. Discussion

Based on Evolutionary Computational approaches presented herein for the identification of multi-dimensional hysteretic systems, it was shown that the obtained equivalent system of differential equations is fairly simple, and at the same time provides reliable estimates with reasonable error. Though the optimization starts off with 9 variables, the discovered equations at the end of the structure optimization by means of GP have only 4 variables, because only the variables will survive during the course of evolution that provide the best fit. It is seen that for the equation of the derivative of the compound restoring force $\dot{r}_i$ at DOF i only the compound restoring force at the same DOF i and the velocity measurements at all DOFs take part in that specific differential equation to attain the best fit for $\dot{r}_i$. The acceptability of the resultant model, despite its simplicity comparing to the complexity of the original system of differential equations, and the physical nature of the studied system, shows the effectiveness of the approach for discovering equivalent computational models for complex multi-dimensional phenomena.

Since only the vibration of lumped masses are measured, without measuring the forces at all interconnecting links associated with the Bouc-Wen elements, the measurements are not complete. Thus the exact original model of Eq. (3) cannot be reconstructed using these measurements by model-free approaches. However, the presented technique of this study is capable of discovering an accurate differential equation for an auxiliary variable that incorporates the vibrations from all other hysteretic links and masses. These differential equations are coupled and together construct the eventual representative model.

7. Advantages and challenges of the proposed technique

7.1 Advantages

In this study, no model was postulated up-front, and the modeling is merely carried out through processing of the provided input and output data. In real world problems, these data are obtained from sensor measurements. Therefore, this method doesn't require extensive prior information about the original underlying system, and as a result the discovered surrogate reduced-order reduced-complexity model can be significantly different from the original system of differential equations while resulting in accurate estimates. Achieving acceptable models through model-free approaches is a great advantage of this technique when dealing with memory-dependent systems. Because these systems do not only depend on the instantaneous values of the variables, their previous states should also be incorporated. Therefore, many conventional non-parametric modeling techniques, such as polynomial approximation, are not able to provide acceptable models for such systems.

Unlike many data-driven approaches, such as neural networks, that yield black-box models, without any insight into the physics of problem, the approach of this paper reveals the computational model of the studied system. On the other hand, if in addition to data, no information about the complex phenomena is available, the discovered model can provide insight into the constitutive properties behind the system associated with the data.

7.2 Challenges

The resultant models highly depend on the provided building blocks in GP, and the embedded basis functions, and not having appropriate basis functions in the pool of building blocks may cause poor results, including too many functions increases the computational cost. Moreover, adding more variables, and consequently more dimensions intensifies the computational cost. The nature of the external excitations also plays a vital role in the modeling process. While, slight perturbations cannot disclose the nature of the hysteretic phenomena to the extent that it is necessary for the modeling, very strong excitation may obscure the hysteretic properties and cause poor generalization capabilities when the system is subjected to less intense excitations.

8. Conclusions

The proposed technique of this paper benefits from advances in the field of Evolutionary Computation to provide high fidelity parsimonious computational models in the form of systems of differential equations that represent multi-dimensional memory-dependent systems, only based on input and output data. This approach employs Genetic Programming to optimize the structures of differential equations, and combines it with Genetic Algorithms to optimize the parameters of formerly discovered structures in a system of coupled differential equations, to result in accurate estimates globally. Thus, basis function and variable selection, dimensionality reduction, and parameter optimization are all preformed in the training phase. A benchmark example is used to assess the effectiveness of the proposed technique. After obtaining the equivalent model for this system based on training excitation, validation is carried out to verify the generalizability of the achieved model in unseen dynamical environments. The validation results show that reasonably accurate responses are achieved though the discovered model is fairly simple compared to the exact formulation of the system, and verify the effectiveness of the presented approach for data-driven modeling of complex multi-dimensional hysteretic system.

References

Alavi, A.H., Gandomi, A.H., Sahab, M.G. and Gandomi, M. (2010), "Multi expression programming: a new approach to formulation of soil classification ", *Eng. Comput.*, **26**(2), 111-118.

Becker, Y., Fei, P. and Lester, A. (2007), "Stock selection: An innovative application of genetic programming methodology", (Eds., Riolo, R., Soule, T. and Worzel, B.), *Genetic Programming Theory and Practice IV, Genetic and Evolutionary Computation*, Springer, USA.

Bolourchi, A. (2014), *Studies into Computational Intelligence Approaches for the Identification of Complex Nonlinear Systems*, PhD Dissertation, University of Southern California, Los Angeles.

Bolourchi, A., Masri, S.F. and Aldraihem, O.J. (2015), "Development and application of computational intelligence approaches for the identification of complex nonlinear systems", *Nonlinear Dynam.*, **79**(2), 765-786, Springer , Netherlands.

Bouc, R. (1967), "Forced vibration of mechanical systems with hysteresis", *Proceedings of the Fourth Conference on Non-linear Oscillation*, Prague, Czechoslovakia.

Chatzi, E.N., Smyth, A.W. and Masri, S.F. (2010), "Experimental application of on-line parametric identification for nonlinear hysteretic systems with model uncertainty", *Struct. Saf.*, **32**(5), 326–337.

Dracopoulos, D. and Effraimidis, D. (2012), *Genetic programming for generalised helicopter hovering control*, *In Genetic Programming, Volume 7244 of Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, Germany.

Gandomi, A.H. and Alavi, A.H. (2011), "Multi-stage genetic programming: a new strategy to nonlinear system modeling", *Inform. Sciences*, **181**(23), 5227-5239.

Gruau, F. (1994), "Automatic definition of modular neural networks", *Adapt. Behav.*, **3**(2), 151-183.

Handley, S. (1995), *Predicting whether or not a nucleic acid sequence is an e. coli promoter region using genetic programming*, Intelligence in Neural and Biological Systems, Symposium on, Herndon, VA, USA.

Hornby, G.S., Lohn, J.D. and Linden, D.S. (2011), "Computer-automated evolution of an x-band antenna for nasa's space technology 5 mission", *Evol. Comput*., **19**(1), 1-23.

Howard, D. and Roberts, S. (2002), *The prediction of journey times on motorways using genetic programming*, Applications of Evolutionary Computing, Berlin, Germany.

Kerschen, G., Worden, K., Vakakis, A. and Golinval, J.C. (2006), "Past, present and future of nonlinear system identification in structural dynamics", *Mech. Syst. Signal. Pr*., **20**(3), 505-592.

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

Li, S., Yu, H. and Suzuki, Y. (2004), "Identification of non-linear hysteretic systems with slip", *Comput. Struct.*, **82**(2), 157-165.

Masri, S., Caffrey, J., Caughey, T., Smyth, A. and Chassiakos, A. (2004), "Identification of the state equation in complex non-linear systems", *Int. J. Nonlinear Mech.*, **39**(7), 1111-1127.

Masri, S., Caffrey, J., Caughey, T., Smyth, A. and Chassiakos, A. (2005), "A general data-based approach for developing reduced-order models of nonlinear mdof systems", *Nonlinear Dynam.*, **39**(1-2), 95-112.

Masri, S. and Caughey, T. (1979), "A nonparametric identification technique for nonlinear dynamic problems", *J. Appl. Mech. - T ASME*, **46**(2), 433-447.

Masri, S., Miller, R., Saud, A. and Caughey, T. (1987), "Identification of nonlinear vibrating structures. ii. applications", *J. Appl. Mech- T ASME.*, **54**(4), 923-950.

Masri, S., Tasbihgoo, F., Caffrey, J., Smyth, A. and Chassiakos, A. (2006a), "Data-based model-free

representation of complex hysteretic mdof systems", *Struct. Control Health.*, **13**(1), 365-387.
Masri, S. F., Ghanem, R., Arrate, F. and Caffrey, J. (2006b), "Stochastic nonparametric models of uncertain hysteretic oscillators", *AIAA J.*, **44**(10), 2319-2330.
Pei, J.S. and Smyth, A.W. (2006a), "New approach to designing multilayer feedforward neural network architecture for modeling nonlinear restoring forces. i: Formulation", *J. Eng. Mech. - ASCE*, **132**(12), 1290–1300.
Pei, J.S. and Smyth, A.W. (2006b), "New approach to designing multilayer feed-forward neural network architecture for modeling nonlinear restoring forces. ii: Applications", *J. Eng. Mech. - ASCE*, **132**(12),1301-1312.
Schmidt, M. and Lipson, H. (2009), "Distilling free-form natural laws from experimental data", *Science*, **324**(5923), 81-85.
Silva, S., Anunciac¸a˜o, O. and Lotz, M. (2011), *A comparison of machine learning methods for the prediction of breast cancer*, Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics, Springer.
Smyth, A., Masri, S., Chassiakos, A. and Caughey, T. (1999), "On-line parametric identification of mdof nonlinear hysteretic systems", *J. Eng. Mech. - ASCE*, **125**(2), 133-142.
Spector, L. and Alpern, A. (1995), "Induction and recapitulation of deep musical structure", *Proceedings of the IJCAI-95 Workshop on Artificial Intelligence and Music*, Montreal, Quebec, Canada.
Tackett, W. (1993), "Genetic programming for feature discovery and image discrimination", *Proceedings of the 5th International Conference on Genetic Algorithms*, San Mateo, CA, USA.
Tasbihgoo, F., Caffrey, J.P. and Masri, S.F. (2007), "Development of data-based model-free representation of non-conservative dissipative systems", *Int. J. Nonlinear Mech.*, **42**(1), 99-117.
Wen, Y.K. (1976), "Method for random vibration of hysteretic systems", *J. Eng. Mech. - ASCE*, **102**(2), 249-263.
Wong, M.L., Leung, K.S. and Cheng, J. (2000), "Discovering knowledge from noisy databases using genetic programming", *J. Am. Soc. Inform. Sci.*, **51**(9), 870 -881.
Zhao, X. and Tan, Y. (2008), "Modeling hysteresis and its inverse model using neural networks based on expanded input space method", *IEEE T. Contr. Syst. T.*, **16**(3), 484-490.