Nonlinear Black-box System Identification through Coevolutionary Algorithms and Radial Basis Function Artificial Neural Networks

Helon Vicente Hultmann Ayala[a,*], Didace Habineza[b], Micky Rakotondrabe[c], Leandro dos Santos Coelho[d,e]

[a] *Department of Mechanical Engineering, Pontifical Catholic University of Rio de Janeiro (PUC-Rio)*
Rua Marques de Sao Vicente, 225, Zip code 22453-900 Rio de Janeiro, RJ, Brazil
[b] *Punch Powertrain nv*
Ondernemerslaan 5429, 3800 Sint-Truiden, Belgium
[c] *Laboratoire Génie de Production (LGP), National School of Engineering in Tarbes (ENIT) / Toulouse INP*
47 Avenue d'Azreix; 65000 Tarbes; France
[d] *Industrial and Systems Engineering Graduate Program, Pontifical Catholic University of Paraná (PUCPR)*
Rua Imaculada Conceição, 1155, Zip code 80215-910 Curitiba, PR, Brazil
[e] *Department of Electrical Engineering, Federal University of Paraná (UFPR)*
Rua Cel. Francisco Heráclito dos Santos, 100, Zip code 81531-980 Curitiba, PR, Brazil

Abstract

The present work deals with the application of coevolutionary algorithms and artificial neural networks to perform input selection and related parameter estimation for nonlinear black-box models in system identification. In order to decouple the resolution of the input selection and parameter estimation, we propose a problem decomposition formulation and solve it by a coevolutionary algorithm strategy. The novel methodology is successfully applied to identify a magnetorheological damper, a continuous polymerization reactor and a piezoelectric robotic micromanipulator. The results show that the method provides valid models in terms of accuracy and statistical properties. The main advantage of the method is the joint input and parameter estimation, towards automating a tedious and error prone procedure with global optimization algorithms.

*Corresponding author. Tel.: +552135271168; Fax: +552135271165
Email addresses: helon@puc-rio.br (Helon Vicente Hultmann Ayala), habinezadidace@gmail.com (Didace Habineza), mrakoton@enit.fr (Micky Rakotondrabe), leandro.coelho@pucpr.br (Leandro dos Santos Coelho)

Keywords: system identification; nonlinear systems; radial basis functions neural networks; coevolutionary algorithms; piezoelectric manipulator

1. Introduction

The field of black-box nonlinear system identification [1] deals with the creation of data-driven mathematical abstractions of dynamic systems with little or no information about its intrinsic properties. The range of its applications spans various domains [2, 3, 4, 5, 6, 7, 8]. The model construction procedure in nonlinear black-box system identification is experimental and involves subjective decisions [9]. After the definition of the experiment protocol, the excitation signal should be devised so that the measured data contains information about the system dynamics. Once data is available, the model structure and related parameters are set so it is possible to evaluate the model quality and assess its adherence to validation metrics. If the model fails to meet the validation criteria, depending on the outputs in terms of model accuracy and residual properties the engineer should reiterate and construct a new model. Being so, the data-driven modeling activity involves subjective decisions and at times tedious and error-prone activities. In the present contribution we try to address one of such issues, focusing on the automatic definition of the inputs of the model.

In this context, Computational Intelligence (CI) [10] paradigms are a source of algorithms in order to alleviate from the engineer some procedures when creating the simulation artefacts. For a review on swarm and evolutionary computing applied to system identification, see [11]. Parameter estimation of infinite impulse response filter models through system identification is made in [12] with a cat swarm algorithm. In [13] the authors introduce a structure identification procedure to evolve fuzzy models using incremental partitioning learning. The cuckoo search algorithm has been applied to optimize an adaptive Hammerstein model in [14]. Solar radiation prediction, which is important in the context of renewable energy dispatch, is made in [15] using different types of Artificial Neural Networks (NNs). In [16] the Covariance Matrix Adaptation

Evolution Strategy (CMA-ES) is implemented to perform neuroevolution of NN controllers for playing the Doom game in visual mode. Evolutionary Algorithms (EAs) are used to discover deep architectures for NNs in [17] and test it in image classification problems, where no human intervention is needed for describing the architecture. Pan evaporation prediction is carried out in [18], showing overall better results for CI methods such as NN and neuro-fuzzy, and in [19] with a genetic fuzzy hybrid method. In [20] the authors use reinforcement learning to generate recurrent NNs for the tasks of image classification and language modeling. Genetic programming has been applied to evolve deep NN structures also for the task of image classification in [21]. In [22] a self-organizing strategy for constructing deep belief networks is proposed and tested for an artificial system, where the architecture is defined by the method iteratively. A variable structure fuzzy wavelet neural network controller has been proposed in [23], which can also be applied to nonlinear adaptive modeling. In [24] the authors use a hybrid training procedure for NNs with Genetic Algorithm (GA) and backpropagation to predict aerosol optical depth.

Bringing CI methods to the context of system identification is however not straightforward. It is necessary to address how these methods adhere to the system identification procedure towards the automation of model generation, which would yield more efficient and cheaper model building activity. Indeed, by automating the tasks in model construction one needs to investigate application specific issues, such as how to avoid to obtain models overly complex which might not be applicable to design feedback control laws. In fact, there are many sources of inspiration which could be drawn in order to build more accurate models with system identification relying more in CI methods.

Coevolutionary Algorithms (CoEAs) [25] rely on the interaction of different populations and their successes in order to evolve the individuals for decision making [26]. Individuals are compared not by their fitness as in traditional EAs, but by their outcomes in the interaction with other populations. Many problems have been tackled by such algorithms by decomposition as results appeared in the literature [27] and also for evolving NNs [28]. CoEAs have been applied in

the context of system identification or function approximation. Reference [29] implements a CoEA with GAs to evolve Takagi-Sugeno (TS) fuzzy models and further extended to soft sensor design in [30] to deal with high-dimensional input space. In [31, 32] the authors implement two populations of a GAs to represent both the lag space and the related parameters of Radial Basis Functions Neural Networks (RBFNNs) for the purpose of time-series forecasting.

In the present paper, we propose the use of a CoEA to perform model-based input selection and related parameter estimation for the special case of system identification and RBFNN models. The coevolutionary approach was able to build accurate and valid models as shown later in the results. This approach is flexible and we may employ different or improved EAs in order to achieve better results. Different types of nonlinear mappings may be used as the metaheuristic paradigm for optimization is very flexible in facing different classes of problems and does not require the cost function to be neither differentiable or convex. The novelties of the present work are highlighted in the following. In the work of Wang and Li [33], the authors propose a CoEA framework for solving the reliability-redundancy problem by dividing the resolution of the problem with continuous and integer decision variables to different EAs. In this contribution, the authors use a Differential Evolution (DE) algorithm for the continuous and an improved Harmony Search (HS) algorithm for the integer variables. The problem of input and parameter estimation can be formulated similarly as in the work from Wang and Li, where in our case the continuous variables represent the NN parameters and the binary variables represent the presence/absence of a lagged input/output in the model. Herein we introduce an improved architecture for the model building problem in system identification based on the concept of CoEA. Specifically, we propose the use of the DE algorithm for optimizing the continuous variables and a recently introduced Adaptive Binary Harmony Search (ABHS) for choosing the inputs of the model. To the best of our knowledge the mentioned CoEA approach has not been applied so far for the purpose of model input selection and related parameter estimation in nonlinear black-box system identification problems. It is important to high-

light that, by using a similar method as in [33] proposed to other domain, it is possible to solve concomitantly the problem of selecting the model inputs and estimating the related model parameters. This is important in the modeling activity as saves the engineer time spent on defining the model architecture and estimating the parameters in an iterative procedure. We test the novel methodology with three case studies: a Magneto-Rheological (MR) damper, a continuous polymerization reactor, and a hysteretic piezoelectric actuator that the authors proposed in a previous work [34]. The results are validated with the tests based on higher-order correlation functions of the residual and the input and the multiple correlation coefficient, showing competitive results when compared to results found in the literature. Moreover, from the results it can be seen that the hysteretic behavior of the pioezoelectric micromanipulator is adequately captured, what motivates the application of the proposed method for building sensorless controllers which are important in this domain. It is also important to highlight that the framework proposed in this paper leverages the automation of the data-driven model building activity, what is desirable with the increasing complexity of systems and the pervasiveness of simulation in various industries for building more accurate models using less human resources. Figure 1 summarizes the approach presented in this paper.

The remainder of the paper is organized as follows. We state the concepts used in the scope of nonlinear black-box system identification input selection and parameter estimation, together with the formulation of the problem in Section 2. The EAs step-by-step procedures are given in Section 3 while in Section 4 the proposed methodology is detailed. The case studies used as references for checking the accuracy of the method are given in Section 5. The results of the proposed methodology are detailed in Section 6 and Section 7 concludes the document, giving general remarks about the study and also future research directions which will be pursued.

2. Nonlinear Black-box System Identification

This section introduces the mathematical formulation of the NN models used for system identification and their validation metrics. Moreover, we state the problem formulation in this context for input selection and parameter estimation which will be the focus of the CoEA methodology proposed for system identification.

In the present work we restrict our attention to Nonlinear AutoRegressive with eXogenous inputs (NARX) models which ar defined as

$$\begin{aligned} y(t) = F[&y(t-1), y(t-2), \ldots, y(t-n_y), \\ &u(t-1), u(t-2), \ldots, u(t-n_u)] + \xi(t). \end{aligned} \tag{1}$$

where $u(t)$ and $y(t)$ are the input and output of the system at a given instant t, considered to be measured. The residual is given by $\xi(t) = y(t) - \hat{y}(t)$, where $\hat{y}(t)$ is the predicted output. Let $F[\cdot]$ be in general a nonlinear function mapping in $\mathbb{R}^{n_\phi} \to \mathbb{R}$, from the n_ϕ model inputs to the predicted output. In this paper we define $F[\cdot]$ as a RBFNN, thus we have

$$\hat{y}(t) = F[\boldsymbol{r}(t)] = \sum_{m=1}^{M} w_m \phi(\boldsymbol{r}(t), \boldsymbol{c}_m, \sigma_m), \tag{2}$$

where $M \in \mathbb{N}^+$ is the number of neurons in the hidden layer and $\boldsymbol{r}(t) \in \mathbb{R}^{n_r}$ is the input vector at a given instant t; $\boldsymbol{c}_m \in \mathbb{R}^{n_r}$, $\sigma_m \in \mathbb{R}^+$, and $w_m \in \mathbb{R}$ are the RBFNN parameters set during training, respectively the center and the width of the m-th hidden node and the output weights. We use the Gaussian activation function for $\phi(\cdot)$ so that we have

$$\phi(\boldsymbol{r}(t), \boldsymbol{c}_m, \sigma_m) = \exp\left[-\frac{1}{2\sigma_m^2} \sum_{i=1}^{n_r} (r_i(t) - c_{m,i})^2\right]. \tag{3}$$

Now we are ready to state the following two problems which will be tackled in the present work.

Problem 1 (Input Selection)**.** *For systems like the ones described in (1), select the model inputs of lagged terms on $u(t)$ and $y(t)$ such that the model represents the physical system accordingly.* □

Problem 2 (Parameter Estimation)**.** *Estimate the related parameters of* $F[\cdot]$ *for the inputs chosen for the model* (1)*, in a way that the model is valid according to some predefined measure.* □

Both problems are critical whenever building a model. It is well known that inadequate inputs lead unavoidably to poor model abstractions. Moreover, often this part is left dependent on the engineer which chooses in a tedious trial and error procedure a suitable set of inputs. Thus, automatic methods to select appropriate inputs which are able to give valid models are highly desirable. Also, the parameter estimation naturally impacts the overall model quality. The engineer builds a model for each new set of inputs of the models chosen, as this part is dependent on input selection. In the following section, we introduce a novel methodology for input selection and parameter estimation based on the concept of coevolution and problem decomposition, which aims at the resolution of Problems 1 and 2 concomitantly.

Let us introduce some model validation metrics in the following. The multiple correlation coefficient (R^2) is calculated as [35]

$$R^2 = \left[1 - \frac{\sum_{t=1}^{N} \xi(t)^2}{\sum_{t=1}^{N} \left(y(t) - \bar{y}\right)^2}\right] \tag{4}$$

where the upper bar denotes the mean value of the the sequence. The R^2 measures the adherence of the model to the measured data, such that $R^2 = 1$ means perfect data reconstruction and $R^2 > 0.9$ may be considered sufficient for many applications [36]. It can be calculated for predictions in (i) One-Step-Ahead (OSA) and (ii) Free-Run simulation (FR). While (i) uses the most recent measured data to perform predictions, in (ii) predictions are made recursively. In general FR-based metrics are more reliable when assessing the model predictive capability and show more easily if the model is valid for longer prediction horizons.

The statistical correlation can also be used for model validation. It indicates whether the estimation procedure was able to explain the data provided for the estimation phase. Usually OSA predictions are taken for calculating the

correlations, as they are more often used as metrics for estimating parameters due to computational issues. According to the set of tests given in [37] for NNs, we have

$$\begin{cases} \phi_{\xi\xi}(\tau) = \delta(\tau), & \\ \phi_{u\xi}(\tau) = 0, & \forall\tau, \\ \phi_{\xi(\xi u)}(\tau) = 0, & \tau \geq 0, \\ \phi_{(u^2)'\xi}(\tau) = 0, & \forall\tau, \\ \phi_{(u^2)'\xi^2}(\tau) = 0, & \forall\tau, \end{cases} \tag{5}$$

where $\delta(\cdot)$ is the Kronecker delta function, $(u^2)'(t) = (u(t))^2 - \bar{u^2}$, $(\xi u) = \xi(t+1)u(t+1)$ and ϕ_{ab} is the normalized cross-correlation function between two sequences $\{a\}$ and $\{b\}$, which is given by

$$\phi_{ab}(\tau) = \frac{\sum_{t=1}^{N-\tau} \left[a(t) - \bar{a}\right] \left[b(t+\tau) - \bar{b}\right]}{\left[\sum_{t=1}^{N} [a(t) - \bar{a}]^2 \sum_{t=1}^{N} [b(t) - \bar{b}]^2\right]^{1/2}}. \tag{6}$$

3. Evolutionary Algorithms

This section briefly outlines the DE and HS algorithms, which compose the CoEA system identification approach. As will be shown, they are used as global search methods in order to jointly define the inputs and the parameters of the black-box models represented by Eq. (1). In the following we depict the step-by-step procedure of both algorithms. All initial solutions are initialized randomly within the search space using an uniform distribution, and the optimization problems are all set as minimization with no loss of generality.

3.1. Differential Evolution

The notation used for the the description of the DE algorithm is given in Table 1. The steps necessary for implementing computationally the algorithm are as follows.

- *Step 1*: Perform mutation;

$$\boldsymbol{v}_i(g) = \boldsymbol{x}_{r_1}(g) + F\left(\boldsymbol{x}_{r_2}(g) - \boldsymbol{x}_{r_3}(g)\right) \tag{7}$$

- *Step 2*: Perform binary crossover;

$$u_{i,j}(g) = \begin{cases} v_{i,j}(g), \text{ if rand} < CR \\ x_{i,j}(g), \text{ otherwise,} \end{cases} \tag{8}$$

where rand is a random number distributed uniformly in the range $[0, 1]$.

- *Step 3*: Evaluation of the newly generated solutions and selection;

$$\boldsymbol{x}_i(g+1) = \begin{cases} \boldsymbol{u}_i(g), \text{ if } f(\boldsymbol{u}_i(g)) < f(\boldsymbol{x}_i(g)) \\ \boldsymbol{x}_i(g), \text{ otherwise,} \end{cases} \tag{9}$$

- *Step 4*: Increment generation counter $g = g + 1$ and check termination criterion. If $g < G$, go to Step 1, otherwise terminate.

The DE algorithm just exposed is used in the present paper for evolving the NN parameters, as will be discussed in the next section.

3.2. Adaptive Binary Harmony Search Algorithm

The ABHS algorithm used in the scope of the present work was proposed in [38]. In this paper, the authors study many adaptive mechanisms to adjust the project parameters of the HS algorithm with binary encoding. In the present work it will be used to denote the evolve the set of candidate lags. The description of the algorithm is given below, while in Table 2 we state the variables used.

- *Step 1*: Adaptively update PAR and HMCR:

$$\text{PAR} = \text{PAR}_0 + \frac{\text{PAR}_f - \text{PAR}_0}{G} \cdot g, \tag{10a}$$

$$\text{HMCR} = \text{HMCR}_0 + \frac{\text{HMCR}_f - \text{HMCR}_0}{G} \cdot g. \tag{10b}$$

- *Step 2*: New harmony improvisation, considering boolean decision variables
  - Harmony memory consideration;

1. Bit selection strategy;

$$h_j^{\text{new}} = \begin{cases} h_{pj}, & \text{if } r_1 < \text{HMCR} \\ \text{round}(r_2), & \text{otherwise} \end{cases} \tag{11}$$

2. Individual selection strategy;

$$h_j^{\text{new}} = \begin{cases} h_{tj}, & \text{if } r_1 < \text{HMCR} \\ \text{round}(r_2), & \text{otherwise} \end{cases} \tag{12}$$

where $t = \text{round}\,(r_3 \cdot \text{HMS})$.

– Pitch adjustment;

$$h_j^{\text{new}} = \begin{cases} h_{bj}, & \text{if } r_1 < \text{PAR} \\ h_j^{\text{new}}, & \text{otherwise} \end{cases} \tag{13}$$

where b is the index of the best harmony found so far by the algorithm.

- *Step 3*: Update HM by replacing the worst stored harmony by $\boldsymbol{h}^{\text{new}}$, if it is better than any of the ones stored.

- *Step 4*: Update generation counter $g = g+1$. Check termination criterion, go to Step 1 if $g < G$ or terminate otherwise.

The ABHS algorithm is used to perform the input selection in the CoEA scheme given in the next section.

4. Proposed Methodology

Now we shall focus to state the proposed methodology based on CoEA to perform input selection and related parameter estimation. We adopt the CoEA paradigm with two populations, namely H and X, in order to decompose the problem of system identification into lag selection and parameter estimation respectively. Being so, H is composed of a set of candidate lags and is evolved by

a binary optimization algorithm while X represents the corresponding RBFNN model parameters (centers and widths) evolved by the continuous optimization algorithm. We shall denote hereafter both populations at the g-th iteration respectively by $H(g)$ and $X(g)$.

The individuals in H and X coevolve during the procedure, and their individuals' fitnesses are calculated according to the ability of the individuals of both populations. It means that if one population performs inappropriately, then unavoidably the performance of the other population will also decrease. In this way, the individuals cooperate as both fitnesses depend on the composed solution formed by them. Each individual of the populations $\boldsymbol{h} \in \{0,1\}^{D_H}$ in H and $\boldsymbol{x} \in \mathbb{R}^{D_X}$ in X is evaluated according to Eq. (14), which represents the objective function to be minimized based on the Mean Squared Error (MSE)

$$f(\boldsymbol{w}) = \text{MSE} = \frac{1}{N}\sum_{t=1}^{N}[\xi(t)]^2 \tag{14}$$

where N represents the amount of data used and $\boldsymbol{w}$ denotes the solution formed by the concatenation of both decision vectors

$$\boldsymbol{w} \triangleq \begin{bmatrix}\boldsymbol{h}^\top & \boldsymbol{x}^\top\end{bmatrix}^\top. \tag{15}$$

We show in Fig. 2 how the individuals of each population are composed. From this illustration we can see that in H the absence or presence of a given lag is indicated by binary variables and in X the individuals encode the corresponding model parameters for the RBFNN, composed by the centers and widths (the output weights are defined by QR factorization [39]).

The dimensions of the decision variables $\boldsymbol{x}$ and $\boldsymbol{h}$ depend on the project parameters given by (i) $n_y^{\max}$ (maximum number of lags in $y(t)$); (ii) $n_u^{\max}$ (maximum number of lags in $u(t)$); and (iii) M (number of neurons of the RBFNN). Being so, the total dimensions of the individuals in H and X are given as defined in Eq. (16):

$$D_H = n_y^{\max} + n_u^{\max}, \tag{16a}$$

$$D_X = (n_y^{\max} + n_u^{\max} + 1)\cdot M. \tag{16b}$$

The objective function aims at the definition of the model parameters based on the MSE of the OSA prediction, which is a quite standard metric for building models. Here, we could in principle employ different types of models but we rather focus on the case of RBFNNs. They are proven to be global approximators [40] being very convenient from the architecture point of view, should the designer select solely the complexity of the RBFNN.

In this approach we adopt the framework for coevolution as described in [33] to solve the optimization problem with the cost function in Eq. (14). In that work, the authors tackle the problem of reliability-redundancy optimization, which is formed by both integer and continuous variables. They adopted a population to handle the integer variables and another to cope with the continuous ones, which coevolve concomitantly in a cooperative fashion. They proposed a modified HS algorithm to deal with the integer variables and a standard DE to manipulate the continuous variables. After half of the overall evolution process, the integer optimization stops and the best candidate is held for the rest of the half of the algorithm, where the continuous variables are further refined.

The framework in [33] is not directly applicable to our problem as H is formed with binary individuals, not integers. Thus, we substituted the modified HS by an ABHS algorithm recently proposed [38], with individual selection strategy. In principle we could employ any other metaheuristic optimization algorithm capable of dealing with binary decision variables. In this case there are many possibilities. ABHS was one of the possible choices, but the best candidate so as to change as little as possible the original framework given by [33] and not increase the number of project parameters. Figures 3 and 4 show the overall procedure and the representation of the individuals for the coevolutionary system identification approach.

Is is important to highlight that we face the problem of input selection and related parameter estimation for system identification as different subproblems which are solved separately in a cooperative fashion. We adopt one population to represent the lags and the other that encodes the parameters of a RBFNN. Being so, we are able to cope with the problem of selecting the input variables

and the parameters of the model concomitantly. In this way the designer is not obliged to set the orders of the model beforehand, what is very convenient specially when e.g. (i) little knowledge about the system is available, (ii) little time is available to spend on trial and error methods for input selection, (iii) there are many inputs and outputs on the system studied, what makes trial and error methods for input selection even more time-consuming.

5. Case Studies

This section is devoted to describe the benchmarks used in the scope of the present work. We describe three dynamic systems given below which will be modeled using their input-output data through the CoEA approach for system identification.

(i) MR Damper: it is a device used for structural vibration control, as it may change its viscosity actively. The data has been used for identification in e.g. [41, 42]. The input is the velocity at the ends of the damper, while the output is the force which it reacts to excitation, that are measured at 200 Hz.

(ii) Continuous polymerization reactor: it describes the free-radical polymerization of MMA (methyl methacrylate) with AIBN (azobisisobutyronitrile) as initiator and toluene as solvent and is supposed that the reaction is made in a jacketed continuous stirred tank reactor. The nonlinear state space model based in first-principles is given in [43]:

$$\dot{z}_1(t) = 10[6 - z_1(t)] - 2.4568\sqrt{z_2(t)}, \tag{17a}$$

$$\dot{z}_2(t) = 80u(t) - 10.1022z_2(t), \tag{17b}$$

$$\dot{z}_3(t) = 0.0024121z_1(t)\sqrt{z_2(t)} + 0.112191z_2(t) - 10z_3(t), \tag{17c}$$

$$\dot{z}_4(t) = 245.978z_1(t)\sqrt{z_2(t)} - 10z_4(t), \tag{17d}$$

$$y(t) = \frac{z_4(t)}{z_3(t)}, \tag{17e}$$

where $u(t)$ is the dimensionless volumetric flow rate of the initiator; $z_1(t)$ is a dimensionless variable representing monomer concentration; $z_2(t)$ represents the initiator concentration; and $y(t)$ is the output of the system, representing

the number-average molecular weight. We generated 50,000 data at 5 Hz, with $u(t) \sim U[0.007, 0.015]$ [44], by solving the nonlinear state equations with the Runge-Kutta method. With the aforementioned input, the range of the output is roughly between 2.6×10^4 and 3.4×10^4 [44]. As in [45], we use 960 data for the purpose of model building and validation. This system has been used for model input selection in [44, 45]. Reference [44] used this simulated system to exemplify an input selection algorithm for nonlinear models based on the false nearest neighbours algorithm using only the input and output data. In [45], the authors apply a novel algorithm to build linear-in-the-parameters models for system identification using genetic programming and orthogonal least squares.

(iii) Piezoelectric micromanipulator: the microactuator, called unimorph, is a free-clamped cantilever composed of two layers: the piezoelectric layer based on lead-zirconate-titanate material (PZT), and a nonpiezoelectric layer based on copper material, see Fig. 5a. The nonpiezoelectric layer is called passive layer. When a voltage u is applied to the piezoelectric layer, it expands or contracts. Due to the constraint between the two layers, this expansion/contraction yields a global deflection (displacement) y of the cantilever (Fig. 5b). Such principle is widely exploited to perform precise positioning [46, 47, 48, 49, 50] because of the nanometric resolution and the high dynamics of the deformation. However, the relation between the voltage u and the output displacement y is known to be hysteretic which drastically affects the final precision. Moreover, this hysteresis was demonstrated to be dynamic, such that the piezoelectric actuator finally exhibits nonlinearities [51]. The experimental setup, presented in Fig. 5c is composed of (i) the piezoelectric actuator with sizes of 15 mm x 2mm x 0.3 mm (length x width x total thickness). The thickness of the piezoelectric layer is 0.2μm while that of the passive layer is 0.1μm; (ii) an optical sensor (LK2420 from Keyence company) which is used to measure the deflection (displacement) of the actuator. Its resolution is tuned to be 40nm and its bandwidth to 5kHz, which are sufficient for the tests in this paper; (iii) a computer (with MATLAB-Simulink) and a dSPACE acquisition board (dS1104) which are used to generate the voltage signal u and to acquire the measurement

y. The sampling time is set to 50μs; and (iv) a voltage amplifier (HV, +-200V) that amplifies the voltage.

6. Results

In the present section, we describe the results obtained by the application of the proposed methodology based on CoEAs for system identification. Three examples are used to test it, namely (i) a MR damper, (ii) a continuous polymerization reactor and (iii) a piezoelectric micromanipulator which were detailed in the Section 5.

The common project parameters used for the simulations in the case studies mentioned above are stated in Table 3, as indicated by [52, 53]. Specifically, for each case study (i), (ii) and (iii) we used $G = 200$, $G = 400$ and $G = 100$ respectively and the quantity of neurons on each RBFNN as $M = 10$ for cases (i) and (ii) and $M = 13$ for (iii), according to the complexity of the system to be modeled. It is important to highlight that the number of generations and model complexity needed for a given system should be set in order to meet model validation requirements. In the present work we established the number of generations so as not to waste computational resources with stalled evaluations after some experimentation. For all case studies we normalized the data in the range [-1,1]. The search range was set for the $c_{m,i}$ in [-4,4] and σ_m in [0.01,20]. The initial population was generated randomly according to a uniform distribution.

Now we focus on the statistical analysis of the final outcome of the proposed methodology. As the search procedures are stochastic, it is necessary to evaluate the outcome of the procedure with different initial conditions. Thus, each case study reported in Section 5 was identified with 30 different initial conditions of the CoEA algorithm. Table 4 shows the statistics of the values obtained by all runs of the CoEA optimization. It is possible to see that the standard deviation is relatively low when considering many different initial conditions. This indicates that the proposed algorithm was able to converge to similar

solutions in the objective space irrespective to the initial conditions, what shows the robustness of the stochastic search approach. Figure 6 shows the progression of the objective function value at each generation considering all runs. The mean value at each generation is plotted, together with a shaded area representing the bounds set with the standard deviation. From this graph it is possible to confirm what was mentioned before for Table 4 that the solutions vary slightly, even though working on completely different initial conditions. We further notice that the algorithm was able to converge after the defined number of generations as expected.

Let us now focus on the analysis of the multiple correlation coefficients obtained by the models, as given by Eq. (4). To this end, we denote by these coefficients obtained in estimation and validation phases respectively by R_e^2 and R_v^2. In Table 5 we depict the metrics in terms of R^2 for the the best models found, after running the procedure with 30 different initial conditions for the algorithm. By best model found we mean that we selected the run with minimal value found for R_v^2 in FR. We can see that in the case of the polymerization reactor and the piezoelectric manipulator, the values of the R^2 metric are close to unity for estimation and validation phases and in OSA and FR, what represents excellent modeling capability. On the other hand, the MR damper case study presented R^2 close to unity in OSA and $R^2 > 0.9$ in FR what should be considered satisfactory as defined by [36] and by our practical appreciation. The excellent accuracy reported in Table 5 is also confirmed in the plots of the measured versus predicted variables. The predictions in OSA and FR can been seen in Figs. 7, 8 and 9 for all case studies, which illustrate R^2 coefficients reported in Table 5 as previously discussed.

Hereafter the results of the solutions presented in Table 5 are detailed. Table 6 shows the set of lags selected by the proposed methodology. In the case of the polymerization reactor, the results reported are in line with previous works which also used the same case study. In [44] the authors apply the false nearest neighbors algorithm to define the order of nonlinear input/output systems, as part of the algorithm herein presented performs. In this work the authors sug-

gest the use of $y(t-1), u(t-1), (t-2)$ for identification. In the results reported in [45], the authors use genetic programming together with orthogonal least squares to discover higher-order polynomial model structures for nonlinear system identification. They have found different sets of lags for various structures, allowing n_y, n_u up to 4 in the overall search procedure. For the most accurate results in average reported, 60% of the trials returned orders as in [44]. Thus, we emphasize that the present methodology, which performs joint input selection and related parameter estimation, presents similar results which are compatible with the ones previously reported in [44, 45]. Specifically, the coevolutionary based approach selected the same lags for $u(t)$ and one additional lag on $y(t)$, out of 10 possibilities which were allowed in each $u(t), y(t)$ if compared to [44] per the CoEA algorithm parameterization (see Table 3). With respect to the MR damper case study, the lags found in the present work are simpler, when compared to the ones found by another method for joint input and parameter estimation method previously reported by the authors in [42], as here we have one less lagged variable in each input and output. It is important to highlight that the set of lags are not unique when considering real-world measured data and the computational burden is much higher in [42] to obtain the models, as for this method to generate a single model we need a couple of hours and the current method instead needs a couple of **seconds**. This is due to the fact that the problem is decomposed by the CoEA formulation, as opposed to the approach in [42].

Now we analyze, for the micromanipulator case study, the ability of the generated model to capture its hysterectic behavior. It is important to highlight that this is another validation requirement for this particular system as this characteristic is known beforehand and important for its real-world application. The path following control algorithms are typically constructed on the basis of hysteresis models such as Bouc-Wen, Preisach, Phase-Preisach combined, Prandtl-Ishlinskii, among other approaches [51]. The goal is to use the information contained in the model about the hysteresis of the system in order to adequately compensate it for improving the accuracy of the regulation. In

Fig. 10 we can see the plot of the output versus the input for the piezoelectric manipulator case study. This is very similar to the one presented in [34]. In this paper the authors have solved this same identification problem testing manually many architectures; in contrast, in the present work we have the selected the inputs and parameters of the model concomitantly, what represents a substantial gain in computational effort. It can thus be seen that the hysteretic behavior has been adequately captured. This reinforces the great opportunity in the application of data-driven modeling tools in order to design the mathematical abstraction for the systems for systems with higher complexity as demonstrated in a recent work which investigated 2-DOF piezoelectric micromanipulators [54] with ad-hoc model structures for NN. Thus, machine learning-based feedback control such as [55] can be directly applied on the basis of the accurate models constructed with less intervention from the engineer, as we have demonstrated.

Finally, Fig. 11 shows the plot of the statistical tests based on the autocorrelation of the residuals and the higher order cross-correlation between the residuals and the system's inputs, given in Eq. (5). In this case it can be seen that the models obtained are also valid in the statistical sense, since the residual analysis points that the dynamics present in the data has been adequately captured. This is important to demonstrate that the models that were built have acquired the dynamics present in the OSA residual information, which was the metric established for the CoEA approach, what confirms the results reported in terms of prediction quality.

7. Conclusion

In the present paper, we showed the application of a novel coevolutionary based methodology for black-box system identification with RBFNN models. We have shown that the present methodology has proven to be suitable for choosing the model orders and related parameters – saving thus time spent on choosing manually the orders of the system or resorting to decoupled methods for input selection and parameter estimation. We chose the HS and DE algo-

rithms tuned with standard parameters to compose the overall model building scheme. The methodology was tested with three nonlinear case studies for system identification, giving accurate and valid results as the R^2 for OSA and FR showed as well as the statistical tests based on correlation functions of the residuals. A piezoelectric manipulator has been used in the scope of the present work, where accurate models have been built with real acquired data, by using the present methodology, and the hysteretic behavior of the system is adequately captured.

In future works, other metaheuristic algorithms may be used in order to evaluate the performance of each combination for the task of coevolutionary system identification. To this end, we may make use of statistical tests in order to evaluate each combination of algorithms that compose the coevolutionary approach [56] or test different approaches [57]. Being so, we may establish what is the best symbiotic relation among algorithms in the CoEA approach. This will be important to devise new research directions on the improvement of each of the EAs that compose the CoEA framework. There is relevant work in the field of evolutionary computation in the creation of novel, more efficient and complex, metaheuristic optimization algorithms based on krill herd [58], beetles [59] and thermal exchange [60] or yet DE improvements [61] to mention a few. It is important that the field of data-driven modeling takes advantage of these new methodologies. Another further study that will deserve attention is the automatic definition of the structure of the nonlinear mapping $F[\cdot]$ and also newer approaches such as [62]. In the case of RBFNNs this regards to the automatic definition of the number of neurons. Some studies have been devoted to the generation of metrics to measure complexity in the scope of NNs applied to system identification [63, 64], which may be applied in studies directed to the extension of the present methodology. A possibility in the specific case of RBFNNs is to adopt the two-steps methodology for training with a clustering technique that defines automatically the number of classes, as the X-means [65] or hybrid approaches [66]. Another idea which should be explored in the context of system identification is the creation of ensembles, which is the use of many

models towards having more precise predictions [67].

Acknowledgment

The authors would like to thank the financial support from CAPES (Coordination for the Improvement of Higher Education Personnel - Brazilian Government) grant no. 88887 . 163347 / 2018-00, the National Council of Scientific and Technologic Development of Brazil – CNPq (Grants number: 303906 / 2015-4-PQ, 404659 / 2016-0-Univ, and 430395/2018-3-Univ) and FA (Fundação Araucária) PRONEX:042/2018 for their financial support.

References

[1] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.-Y. Glorennec, H. Hjalmarsson, A. Juditsky, Nonlinear black-box modeling in system identification: a unified overview, Automatica 31 (12) (1995) 1691–1724.

[2] V. Valdivia, A. Lazaro, A. Barrado, P. Zumel, C. Fernandez, M. Sanz, Black-box modeling of three-phase voltage source inverters for system-level analysis, IEEE Transactions on Industrial Electronics 59 (9) (2012) 3648–3662.

[3] V. Valdivia, A. Barrado, A. Lazaro, M. Sanz, D. Lopez del Moral, C. Raga, Black-box behavioral modeling and identification of dc–dc converters with input current control for fuel cell power conditioning, IEEE Transactions on Industrial Electronics 61 (4) (2014) 1891–1903.

[4] Y. Kim, R. Mallick, S. Bhowmick, B.-L. Chen, Nonlinear system identification of large-scale smart pavement systems, Expert Systems with Applications 40 (9) (2013) 3551 – 3560.

[5] F. da Costa Lopes, E. Watanabe, L. Rolim, A control-oriented model of a PEM fuel cell stack based on NARX and NOE neural networks, IEEE Transactions on Industrial Electronics 62 (8) (2015) 5155–5163.

[6] P. Gil, F. Santos, L. Palma, A. Cardoso, Recursive subspace system identification for parametric fault detection in nonlinear systems, Applied Soft Computing 37 (2015) 444 – 455.

[7] T. A. Tutunji, Parametric system identification using neural networks, Applied Soft Computing 47 (2016) 251 – 261.

[8] H. Garrido, O. Curadelli, D. Ambrosini, A heuristic approach to output-only system identification under transient excitation, Expert Systems with Applications 68 (2017) 11 – 20.

[9] P. M. Nørgård, O. Ravn, N. K. Poulsen, L. K. Hansen, Neural networks for modelling and control of dynamic systems: a practitioner's handbook, Springer-Verlag, London, 2000.

[10] A. P. Engelbrecht, Computational Intelligence: An Introduction, 2nd Edition, Wiley, Chichester, UK, 2007.

[11] A. Gotmare, S. S. Bhattacharjee, R. Patidar, N. V. George, Swarm and evolutionary computing algorithms for system identification and filter design: A comprehensive review, Swarm and Evolutionary Computation 32 (2017) 68 – 84.

[12] G. Panda, P. M. Pradhan, B. Majhi, IIR system identification using cat swarm optimization, Expert Systems with Applications 38 (10) (2011) 12671 – 12683.

[13] H. H. Y. Sa'ad, N. A. M. Isa, M. M. Ahmed, A. H. Y. Sa'd, A robust structure identification method for evolving fuzzy system, Expert Systems with Applications 93 (2018) 267 – 282.

[14] A. Gotmare, R. Patidar, N. V. George, Nonlinear system identification using a cuckoo search optimized adaptive hammerstein model, Expert Systems with Applications 42 (5) (2015) 2538 – 2546.

[15] L. Wang, O. Kisi, M. Zounemat-Kermani, G. A. Salazar, Z. Zhu, W. Gong, Solar radiation prediction using different techniques: model evaluation and comparison, Renewable and Sustainable Energy Reviews 61 (2016) 384 – 397.

[16] S. Alvernaz, J. Togelius, Autoencoder-augmented neuroevolution for visual doom playing, in: IEEE Conference on Computational Intelligence and Games, New York, NY, USA, 2017.

[17] E. Real, S. Moore, A. Selle, S. Saxena, Y. L. Suematsu, J. Tan, Q. V. Le, A. Kurakin, Large-scale evolution of image classifiers, in: D. Precup, Y. W. Teh (Eds.), Proceedings of the 34th International Conference on Machine Learning, Vol. 70 of Proceedings of Machine Learning Research, PMLR, International Convention Centre, Sydney, Australia, 2017, pp. 2902–2911.

[18] L. Wang, O. Kisi, M. Zounemat-Kermani, H. Li, Pan evaporation modeling using six different heuristic computing methods in different climates of china, Journal of Hydrology 544 (2017) 407 – 427.

[19] L. Wang, O. Kisi, B. Hu, M. Bilal, M. Zounemat-Kermani, H. Li, Evaporation modelling using different machine learning techniques, International Journal of Climatology 37 (S1) (2017) 1076–1092.

[20] Q. L. Barret Zoph, Neural architecture search with reinforcement learning, in: International Conference on Learning Representations, Toulon, France, 2017.

[21] M. Suganuma, S. Shirakawa, T. Nagao, A genetic programming approach to designing convolutional neural network architectures, in: Proceedings of the Genetic and Evolutionary Computation Conference, GECCO '17, ACM, Berlin, Germany, 2017, pp. 497–504.

[22] J. Qiao, G. Wang, X. Li, W. Li, A self-organizing deep belief network for nonlinear system modeling, Applied Soft Computing 65 (2018) 170 – 183.

[23] Y. Solgi, S. Ganjefar, Variable structure fuzzy wavelet neural network controller for complex nonlinear systems, Applied Soft Computing 64 (2018) 674 – 685.

[24] W. Qin, L. Wang, A. Lin, M. Zhang, M. Bilal, Improving the estimation of daily aerosol optical depth and aerosol radiative effect using an optimized artificial neural network, Remote Sensing 10 (7) (2018).

[25] M. A. Potter, K. A. De Jong, Cooperative coevolution: An architecture for evolving coadapted subcomponents, Evolutionary Computation 8 (1) (2000) 1–29.

[26] E. Popovici, A. Bucci, R. Wiegand, E. D. De Jong, Coevolutionary principles, in: G. Rozenberg, T. Bäck, J. N. Kok (Eds.), Handbook of Natural Computing, Springer, Berlin, Germany, 2012, pp. 987–1033.

[27] C. Wang, J.-H. Gao, A differential evolution algorithm with cooperative coevolutionary selection operation for high-dimensional optimization, Optimization Letters 8 (2) (2014) 477–492.

[28] F. Gomez, J. Schmidhuber, R. Miikkulainen, Accelerated neural evolution through cooperatively coevolved synapses, The Journal of Machine Learning Research 9 (2008) 937–965.

[29] M. R. Delgado, F. V. Zuben, F. Gomide, Coevolutionary genetic fuzzy systems: a hierarchical collaborative approach, Fuzzy Sets and Systems 141 (1) (2004) 89–106.

[30] M. R. Delgado, E. Y. Nagai, L. V. R. de Arruda, A neuro-coevolutionary genetic fuzzy system to design soft sensors, Soft Computing 13 (5) (2009) 481–495.

[31] E. Parras-Gutierrez, V. Rivas, M. Garcia-Arenas, M. del Jesus, Short, medium and long term forecasting of time series using the L-Co-R algorithm, Neurocomputing 128 (2014) 433 – 446.

[32] E. Parras-Gutierrez, V. M. Rivas, J. J. Merelo, A Radial Basis Function Neural Network-Based Coevolutionary Algorithm for Short-Term to Long-Term Time Series Forecasting, Springer International Publishing, Vilamoura, Portugal, 2016, pp. 121–136.

[33] L. Wang, L.-P. Li, A coevolutionary differential evolution with harmony search for reliability-redundancy optimization, Expert Systems with Applications 39 (5) (2012) 5271–5278.

[34] H. V. H. Ayala, D. Habineza, M. Rakotondrabe, C. E. Klein, L. S. Coelho, Nonlinear black-box system identification through neural networks of a hysteretic piezoelectric robotic micromanipulator, in: 17th IFAC Symposium on System Identification, Beijing, China, 2015.

[35] R. Haber, H. Unbehauen, Structure identification of nonlinear dynamic systems–a survey on input/output approaches, Automatica 26 (4) (1990) 651–677.

[36] B. Schaible, H. Xie, Y.-C. Lee, Fuzzy logic models for ranking process effects, IEEE Transactions on Fuzzy Systems 5 (4) (1997) 545–556.

[37] S. Billings, H. Jamaluddin, S. Chen, Properties of neural networks with applications to modelling non-linear dynamical systems, International Journal of Control 55 (1) (1992) 193–224.

[38] L. Wang, R. Yang, Y. Xu, Q. Niu, P. M. Pardalos, M. Fei, An improved adaptive binary harmony search algorithm, Information Sciences 232 (2013) 58–87.

[39] G. H. Golub, C. F. Van Loan, Matrix computations, 4th Edition, JHU Press, Baltimore, USA, 2013.

[40] J. Park, I. W. Sandberg, Universal approximation using radial-basis-function networks, Neural Computation 3 (2) (1991) 246–257.

[41] J. Wang, A. Sano, T. Chen, B. Huang, Identification of Hammerstein systems without explicit parameterisation of non-linearity, International Journal of Control 82 (5) (2009) 937–952.

[42] H. V. H. Ayala, L. dos Santos Coelho, Cascaded evolutionary algorithm for nonlinear system identification based on correlation functions and radial basis functions neural networks, Mechanical Systems and Signal Processing 68 (2016) 378–393.

[43] F. J. Doyle III, B. A. Ogunnaike, R. K. Pearson, Nonlinear model-based control using second-order Volterra models, Automatica 31 (5) (1995) 697–714.

[44] C. Rhodes, M. Morari, Determining the model order of nonlinear input/output systems, AIChE Journal 44 (1) (1998) 151–163.

[45] J. Madár, J. Abonyi, F. Szeifert, Genetic programming for the identification of nonlinear input-output models, Industrial & Engineering Chemistry Research 44 (9) (2005) 3178–3186.

[46] J. Agnus, N. Chaillet, C. Clévy, S. Dembél´e, M. Gauthier, Y. Haddab, G. Laurent, P. Lutz, N. Piat, K. Rabenorosoa, M. Rakotondrabe, B. Tamadazte, Robotic microassembly and micromanipulation at femto-st, Journal of Micro-Bio Robotics 8 (2) (2013) 91–106.

[47] G. Binnig, D. P. E. Smith, Single-tube three-dimensional scanner for scanning tunneling microscopy, Review of Scientific Instruments 57 (8) (1986) 1688–1689.

[48] S. Lescano, D. Zlatanov, M. Rakotondrabe, N. Andreff, Kinematic analysis of a meso-scale parallel robot for laser phonomicrosurgery, in: A. Kecskeméthy, F. Geu Flores (Eds.), Interdisciplinary Applications of Kinematics, Vol. 26 of Mechanisms and Machine Science, Springer International Publishing, 2015, pp. 127–135.

[49] S. Mitsuhashi, K. Wakamatsu, Y. Aihara, N. Okihara, Relay using multilayer piezoelectric actuator, Japanese Journal of Applied Physics 24 (S3) (1985) 190–192.

[50] S. Devasia, E. Eleftheriou, S. Moheimani, A survey of control issues in nanopositioning, IEEE Transactions on Control Systems Technology 15 (5) (2007) 802–823.

[51] M. Rakotondrabe, Multivariable classical prandtl–ishlinskii hysteresis modeling and compensation and sensorless control of a nonlinear 2-dof piezoactuator, Nonlinear Dynamics 89 (1) (2017) 481–499.

[52] Z. W. Geem, Optimal cost design of water distribution networks using harmony search, Engineering Optimization 38 (3) (2006) 259–277.

[53] D. Simon, Evolutionary optimization algorithms: Biologically-Inspired and Population-Based Approaches to Computer Intelligence, John Wiley & Sons, Hoboken, USA, 2013.

[54] H. V. H. Ayala, M. Rakotondrabe, L. S. Coelho, Modeling of a 2-dof piezoelectric micromanipulator at high frequency rates through nonlinear black-box system identification, in: Proceedings of the American Control Conferece, ACC '18, Milwaukee, USA, 2018.

[55] J. de Jesús Rubio, Discrete time control based in neural networks for pendulums, Applied Soft Computing 68 (2018) 821 – 832.

[56] J. Derrac, S. García, D. Molina, F. Herrera, A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms, Swarm and Evolutionary Computation 1 (1) (2011) 3–18.

[57] W. Puchalsky, G. T. Ribeiro, C. P. da Veiga, R. Z. Freire, L. dos Santos Coelho, Agribusiness time series forecasting using wavelet neural networks and metaheuristic optimization: An analysis of the soybean sack

price and perishable products demand, International Journal of Production Economics 203 (2018) 174 – 189.

[58] A. H. Gandomi, A. H. Alavi, Krill herd: A new bio-inspired optimization algorithm, Communications in Nonlinear Science and Numerical Simulation 17 (12) (2012) 4831 – 4845.

[59] N. A. Kallioras, N. D. Lagaros, D. N. Avtzis, Pity beetle algorithm – a new metaheuristic inspired by the behavior of bark beetles, Advances in Engineering Software 121 (2018) 147 – 166.

[60] A. Kaveh, A. Dadras, A novel meta-heuristic optimization algorithm: Thermal exchange optimization, Advances in Engineering Software 110 (2017) 69 – 84.

[61] M. Tsili, E. I. Amoiralis, J. V. Leite, S. R. Moreno, L. dos Santos Coelho, A novel multiobjective lognormal-beta differential evolution approach for the transformer design optimization, Engineering Computations 35 (2) (2018) 955–978.

[62] B. Zhang, S. Billings, Identification of continuous-time nonlinear systems: The nonlinear difference equation with moving average noise (NDEMA) framework, Mechanical Systems and Signal Processing 60-61 (2015) 810–835.

[63] H. M. R. Ugalde, J.-C. Carmona, V. M. Alvarado, J. Reyes-Reyes, Neural network design and model reduction approach for black box nonlinear system identification with reduced number of parameters, Neurocomputing 101 (2013) 170 – 180.

[64] H. M. R. Ugalde, J.-C. Carmona, J. Reyes-Reyes, V. M. Alvarado, J. Mantilla, Computational cost improvement of neural network models in black box nonlinear system identification, Neurocomputing 166 (2015) 96 – 108.

[65] D. Pelleg, A. Moore, X-means: Extending K-means with efficient estimation of the number of clusters, in: 17th International Conference on Machine Learning, Stanford, CA, USA, 2000.

[66] B. Subudhi, D. Jena, Nonlinear system identification using memetic differential evolution trained neural networks, Neurocomputing 74 (10) (2011) 1696 – 1709.

[67] G. T. Ribeiro, V. C. Mariani, L. dos Santos Coelho, Enhanced ensemble structures using wavelet neural networks applied to short-term load forecasting, Engineering Applications of Artificial Intelligence 82 (2019) 272 – 281.

Figures

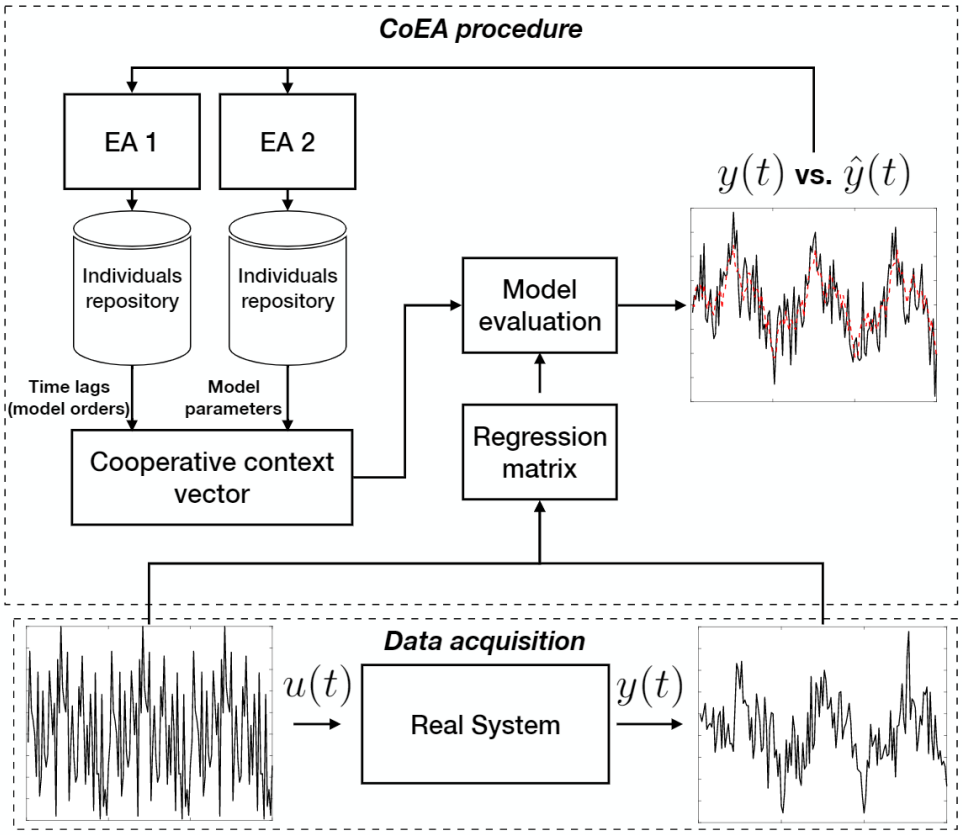

Figure 1: Summary of the proposed methodology. The CoEA estimates the orders of the model and the model parameters in a cooperative fashion.The response of the model is compared to the one measured, and this information is used to evaluate the individuals in each population.

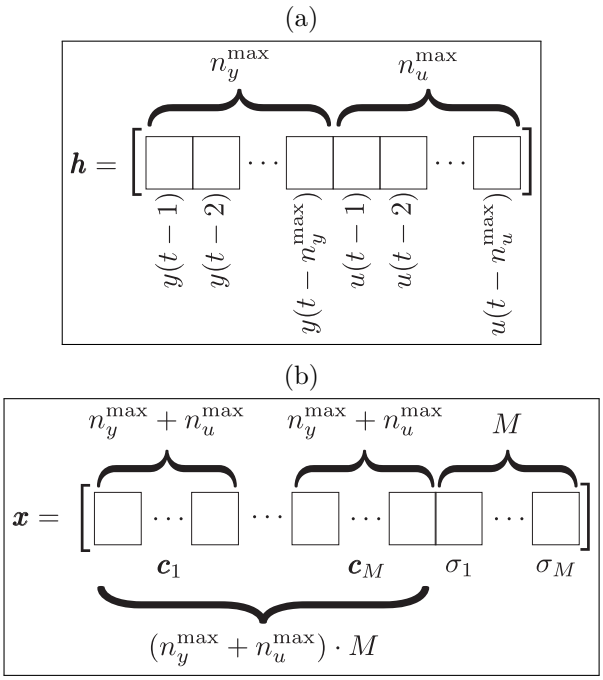

Figure 2: Encoding of the individuals (a) $\boldsymbol{h}$ and (b) $\boldsymbol{x}$ for the proposed coevolutionary approach.

Algorithm: Coevolutionary System Identification

begin

Initialize $H(0)$ and $X(0)$ randomly according to an uniform distribution inside the search space;

Evaluate $H(0)$ and $X(0)$ by randomly selecting one individual in the other population according to an uniform distribution;

for $g \leftarrow 1$ **to** G **do**

Update HMCR and PAR;

foreach $\boldsymbol{x}_i(g) \in X(g)$ **do**

Generate a trial vector $\boldsymbol{v}_i(g)$ (DE);

if $g < \dfrac{G}{2}$ **then**

Generate a new harmony $\boldsymbol{h}$ based on $H(g)$;

Evaluate the solution $\boldsymbol{w} = \left[\boldsymbol{h}^\top \quad \boldsymbol{v}_i(g)^\top\right]^\top$;

Replace the worst individual in $H(g)$;

else

Select the best individual $\boldsymbol{h}_b \in H(g)$;

Evaluate the solution $\boldsymbol{w} = \left[\boldsymbol{h}_b^\top \quad \boldsymbol{v}_i(g)^\top\right]^\top$;

end

end

Update $X(g)$ using greedy selection (DE);

end

end

Figure 3: Pseudocode for the coevolutionary system identification approach based on HS and DE.

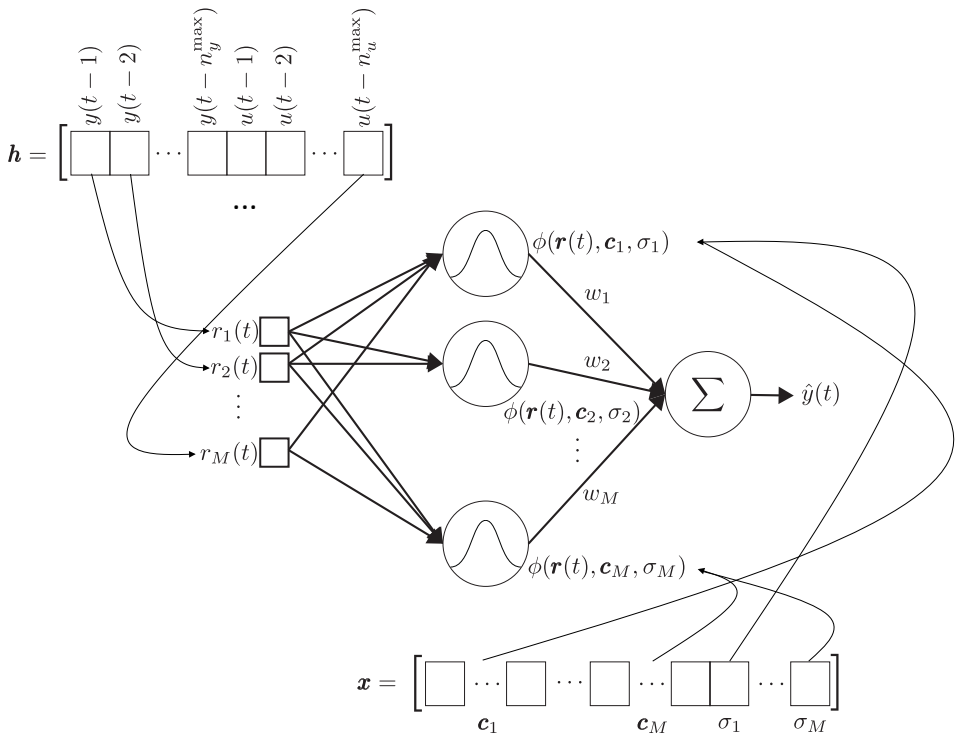

Figure 4: Coevolutionary approach adopted for system identification.

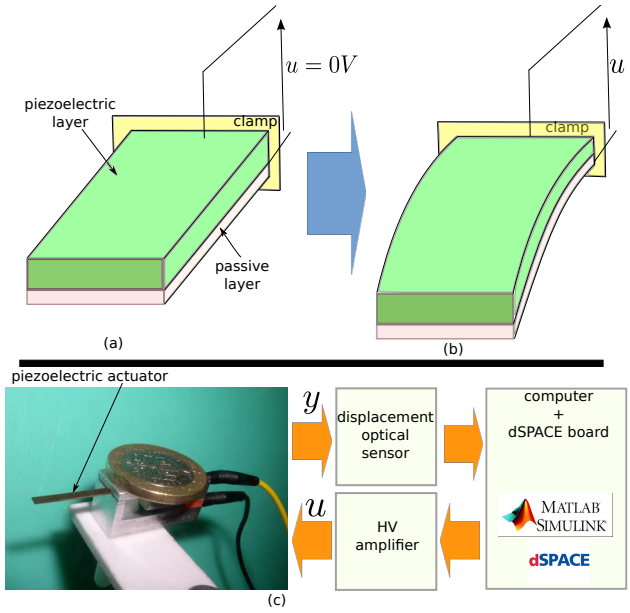

Figure 5: The experimental setup. (a) and (b): principle of functioning of the piezoelectric actuator. (c): the experimental setup diagram.

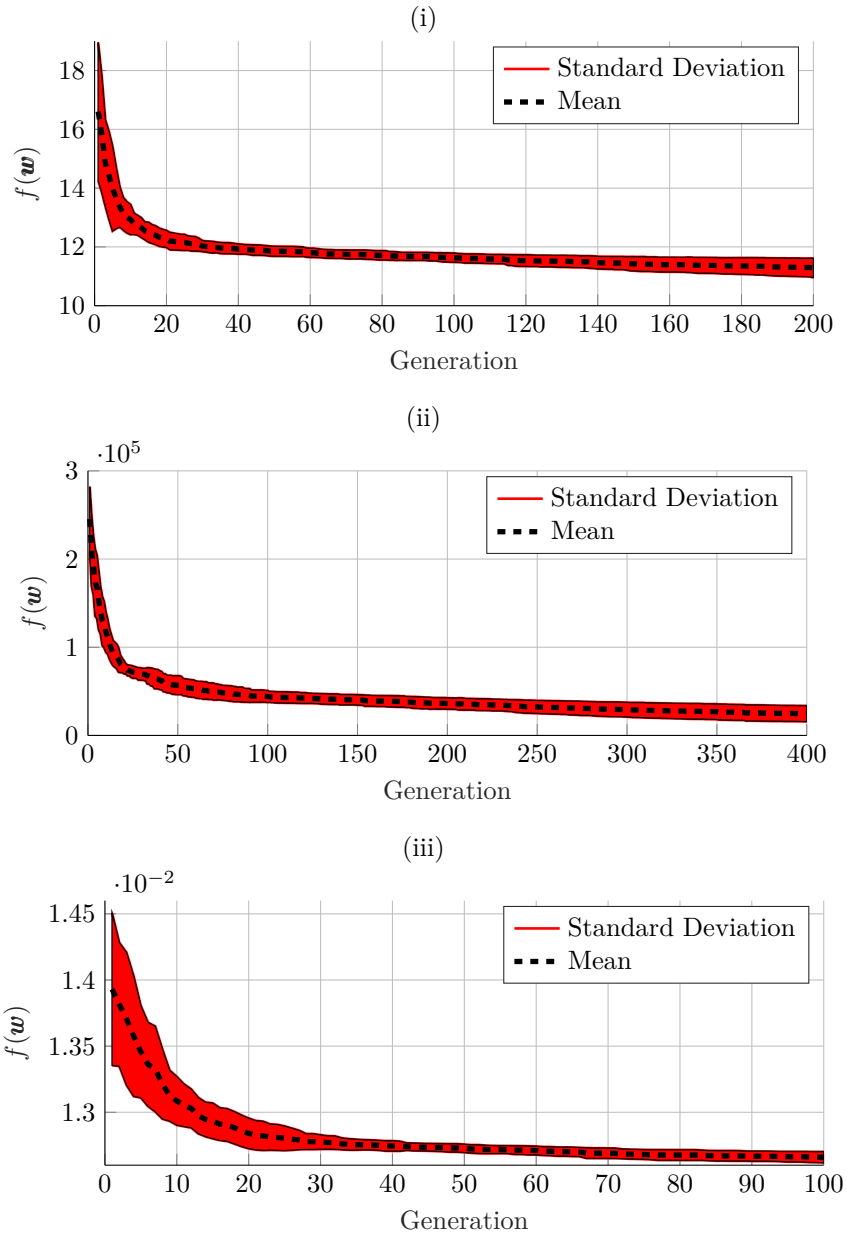

Figure 6: Evolution at each generation of the mean and standard deviation of the objective function for the (i) MR damper , (ii) polymerization reactor and (iii) piezoelectric actuator. In dotted lines we see the mean of the objective function value measured at each iteration of the algorithm, while in red we see area covered by the standard deviation around the mean value.

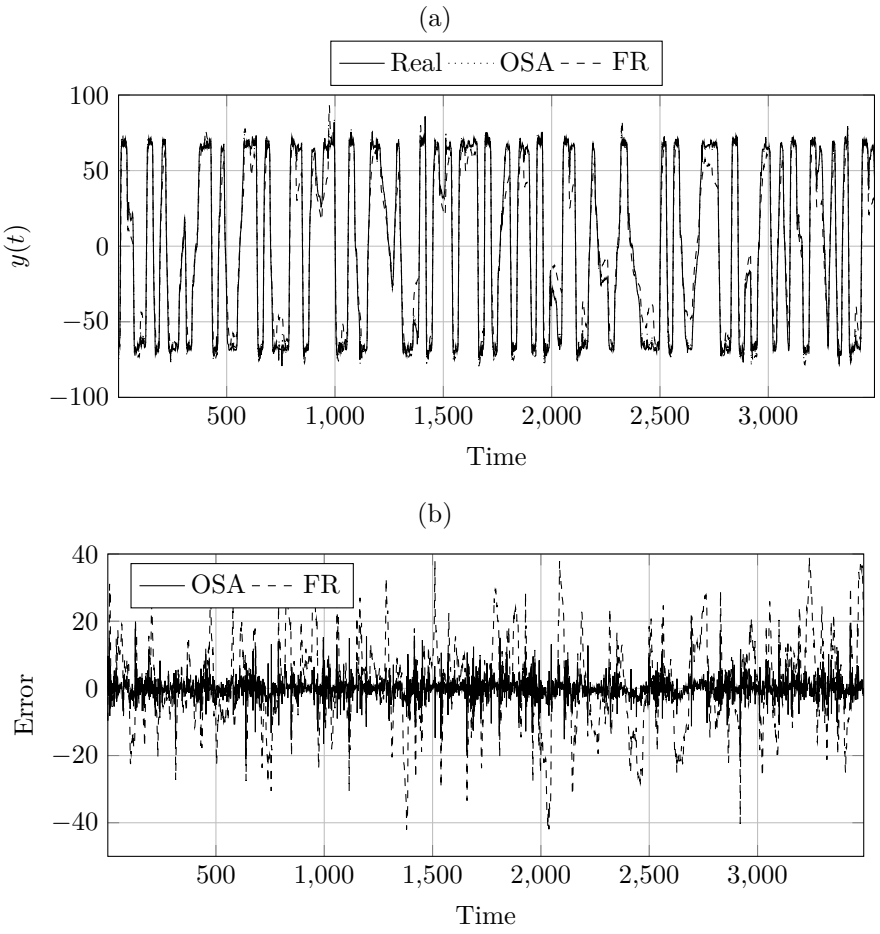

Figure 7: Predictions and errors for the MR damper case study.

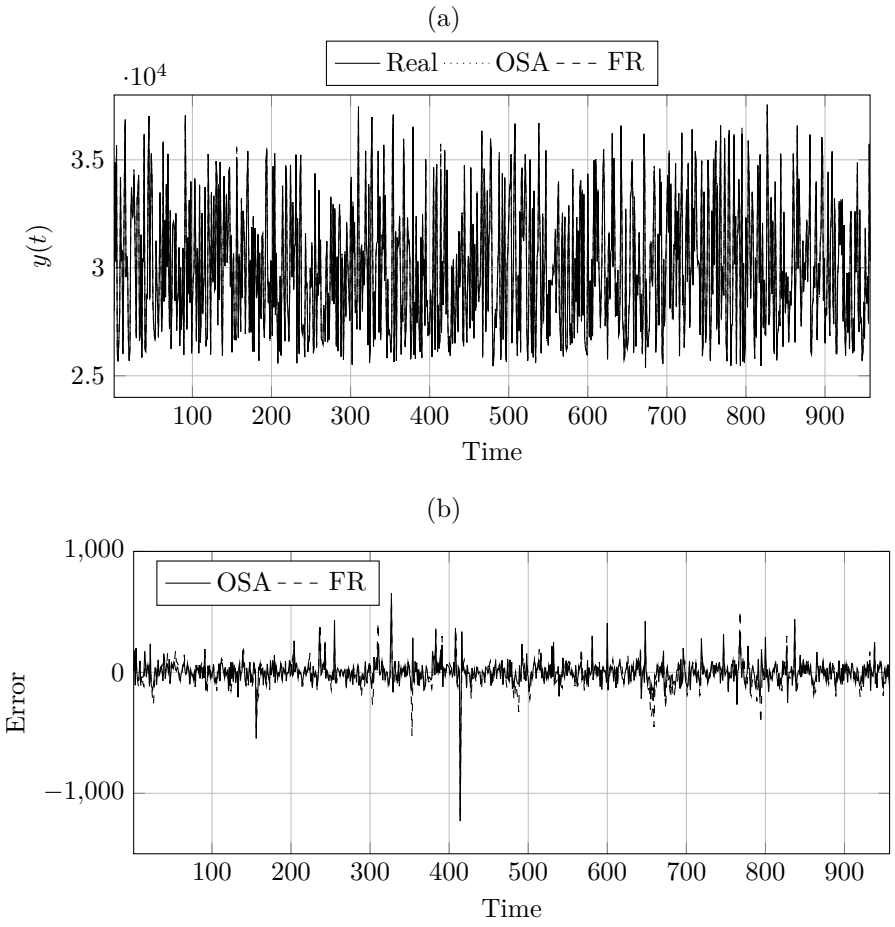

Figure 8: Predictions and errors for the polymerization reactor case study.

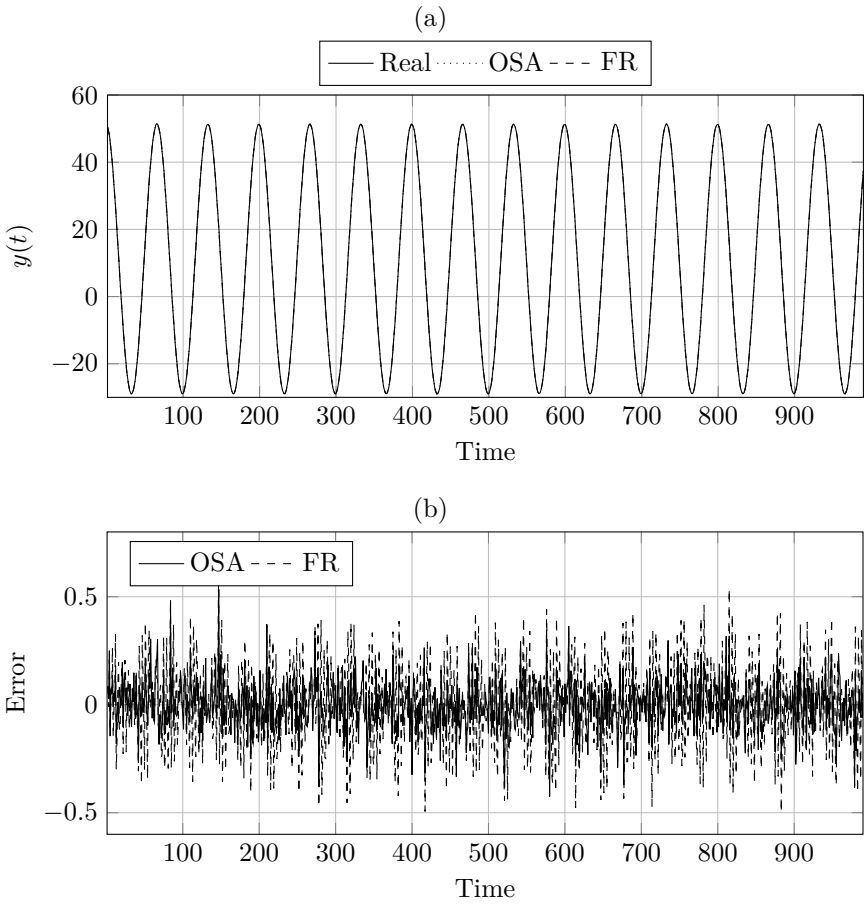

Figure 9: Predictions and errors for the piezoelectric actuator case study.

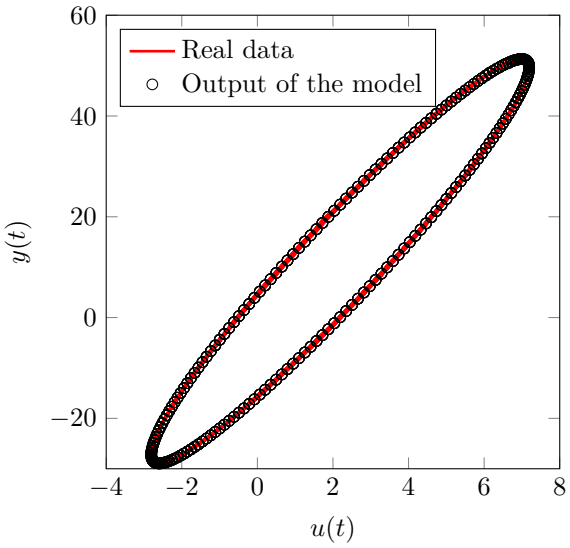

Figure 10: Plot of the $u(t)$ versus the real output and the free simulation result, for the piezoelectric actuator case study. Note that the hysteretic behavior has been adequately captured.

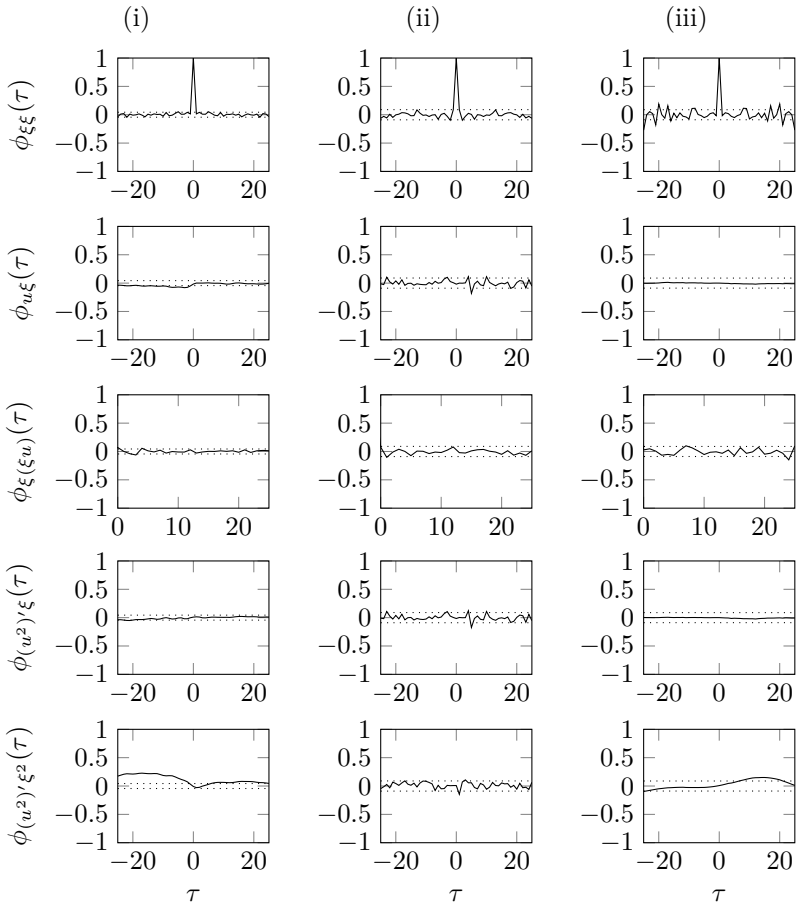

Figure 11: Tests based on the correlation of the residuals and inputs for the (i) MR damper case study, (ii) the continuous polymerization reactor and (iii) the piezoelectric actuator.

Tables

Table 1: Description of the nomenclature for the variables used in the DE optimization algorithm.

Symbol	**Description**
NP	Number of individuals in a population
D	Problem dimension
g	Generation counter
G	Total number of generations
F	Scale factor, in $[0, 1]$
CR	Crossover probability), in $[0, 1]$
$\boldsymbol{x}_i(g)$	i-th individual vector
$\boldsymbol{v}_i(g)$	i-th mutant vector
$\boldsymbol{u}_i(g)$	i-th trial vector
r_1, r_2, r_3	Uniformly random integers, mutually different with the target vector index, in the range $[1, NP]$

Table 2: Description of the nomenclature for the variables used in the HS optimization algorithm.

Symbol	**Description**
HMCR_0, HMCR_f	Initial and final harmony memory considering rate, respectively
PAR_0, PAR_f	Initial and final pitch adjusting rate, respectively
bw_j	Pitch bandwidth for the j-th dimension
HM	Harmony matrix
HMS	Number of solutions HM
$\boldsymbol{h}_i$	$i-$th binary harmony stored in HM
$\boldsymbol{h}^{\mathrm{new}}$	New solution generated according to ABHS strategies
r_1, r_2, r_3	Random numbers uniformly distributed in the range $[0, 1]$

Table 3: Parameters used in the CoEA for system identification.

Parameter description	Value
Number of candidate lags on $y(t)$ and $u(t)$	$n_y^{\max} = n_u^{\max} = 10$
Population size (both H and X)	20
DE type	DE/rand/1
Crossover probability	$CR = 0.8$
Scale factor	$F = 0.6$
Harmony memory consideration rate	$\begin{cases} \text{HMCR}_0 = 0.8 \\ \text{HMCR}_f = 1 \end{cases}$
Pitch adjusting rate	$\begin{cases} \text{PAR}_0 = 0.1 \\ \text{PAR}_f = 0.9 \end{cases}$

Table 4: Minimum (min.), maximum (max.), mean and standard deviation (std. dev.) of the results in terms of $f(\boldsymbol{w})$ for the case studies analyzed.

Case study	Min.	Max.	Mean	Std. dev.
MR damper	10.5049	11.8069	11.2856	0.3273
Pol. reactor ($\times 10^4$)	0.9806	4.7628	2.4681	0.8897
Piezoelectric actuator	0.012570	0.01273	0.01266	4.1034E-05

Table 5: Multiple correlation coefficients for the case studies analysed in OSA and FR, for estimation (est.) and validation (val.) phases.

	OSA		FR	
Case study	R^2 (est.)	R^2 (val.)	R^2 (est.)	R^2 (val.)
MR damper	0.9968	0.9971	0.9716	0.9591
Polymerization reactor	0.9989	0.9993	0.9987	0.9992
Piezoelectric actuator	0.99998	0.99998	0.99995	0.99995

Table 6: Best regressors found by the coevolutionary based methodology.

Case study	Regressors chosen	
MR damper		$u(t-1)$
	$y(t-1)$	$u(t-2)$
	$y(t-2)$	$u(t-6)$
		$u(t-7)$
Continuous polymerization reactor	$y(t-1)$	$u(t-1)$
	$y(t-2)$	$u(t-2)$
Piezoelectric actuator	$y(t-1)$	$u(t-1)$
	$y(t-2)$	$u(t-2)$
	$y(t-3)$	$u(t-6)$
	$y(t-4)$	$u(t-8)$
	$y(t-8)$	$u(t-10)$