Multi-criteria decision and multi-objective optimization for constructing and selecting models for systems identification

J. Hernández-Riveros & A. Arboleda-Gómez
Facultad de Minas, Universidad Nacional de Colombia – Sede Medellín, Colombia

Abstract

An alternative form for the identification of dynamic systems with the application of multi-objective optimization concepts, through the evolutionary algorithm MAGO is presented. A computational tool using operational data of a SISO system has been designed to automatically perform the construction and selection of the best model representing it. After a data acquisition, strategies for the system identification by parametric modelling are developed. The application on the fitness function of appropriate criteria to choose models representing the system is also studied. Different models (ARX, ARMAX, and OE) are built and compared. The models obtained, by evolution, provide better fit and final prediction error regarding that chosen by an expert. The computational effort is low considering that the proposed method is more effective on identification of dynamic systems. Applying this evolutionary method to more complex systems such as MISO, MIMO, and non-linear is proposed as future work.

Keywords: computational and experimental methods, system identification, evolutionary computation, multi-criteria decision, multi-objective optimization.

1 Introduction

Representing the dynamics of a system by mathematical models from measured data is the objective of the system identification. Parametric models as ARX, ARMAX, OE, and others are employed for this purpose. Finding a model with an appropriate structure representing a system is a work that represents a considerable time demand. Typically, after applying system identification

WIT Transactions on Modelling and Simulation, Vol 55,
www.witpress.com, ISSN 1743-355X (on-line)
doi:10.2495/CMEM130031

methods, this leads to having several candidate models representing the system. The task of selecting models increases its complexity when the number of models grows significantly. To solve this matter, using an evolutionary algorithm a computational tool has been developed. This tool, from the system's input and output data, builds, evaluates, and selects automatically, quickly, and efficiently the most appropriate identification model to represent a given problem. This paper begins with an introduction to systems identification, multi-objective optimization, and evolutionary algorithms. It continues with the approach of their integration in a tool for constructing and selecting parametric models for dynamic systems. Results of two cases of study, one a known academic laboratory and the other a standard problem, are presented. It ends with conclusions and future work.

2 Systems identification

System identification can build mathematical models that represent roughly the dynamics of a system from measured data of input and output variables [2]. To define the structure of the problem may propose a new one or apply some known ones. Among the commonly used statistical models to represent the dynamics of a system are ARX, ARMAX, OE, BJ ARIMAX, and others [3]. For the study cases in this article the first three mentioned parametric models are applied.

2.1 ARX models

Present results in a linear regression. The ARX structure is a linear difference equation of the kind of equation (1).

$$\begin{aligned} y(t) + \ a_1 y(t-1) + \cdots + a_{na} y(t-na) \\ = b_1 u(t-nk) + \cdots + b_{nb} u(t-nk-nb+1) \end{aligned} \tag{1}$$

The model can be written as

$$A(q)y(t) = B(q)u(t-nk) + e(t) \tag{2}$$

2.2 ARMAX models

Possess several improvements in relation to the ARX model, such as a different modeling of the noise. The ARMAX model is given by

$$\begin{aligned} y(t) + \ a_1(t-1) + \cdots + a_{na} y(t-na) = \ b_1 u(t-nk) + \cdots \\ + b_{nb} u(t-nk-nb+1 + e(t) + c_1 e(t-1) + \cdots + \ e_{nc} e(t-nc) \end{aligned} \tag{3}$$

The equation is rewritten as

$$A(q)y(t) = B(q)u(t-nk) + C(q)e(t) \tag{4}$$

where q is the lag operator and polynomials A (q), B (q), and C (q) are given by

$$\begin{aligned} A(q) &= 1 + a_1 q^{-1} + \ a_2 q^{-2} + \cdots + \ a_{na} q^{-na} \\ B(q) &= b_1 + b_2 q^{-1} + \ b_2 q^{-2} + \cdots + \ b_{nb} q^{-nb+1} \\ C(q) &= 1 + \ C_1 q^{-1} + \cdots + C_{nc} q^{-nc} \end{aligned} \tag{5}$$

2.3 OE models

Similar to the ARMAX model structure, use the Likelihood method to estimate the parameters of the polynomials. The model structure OE (Output Error) is given by

$$y(t) = \frac{B(q)}{F(q)} u(t - nk) + e(t) \tag{6}$$

where the polynomials B (q) and F (q) are given by

$$B(q) = b_1 + b_2 q^{-1} + \; b_2 q^{-2} + \cdots + \; b_{nb} q^{-nb+1} \tag{7}$$
$$F(q) = 1 + f_1\, q^{-1} + \cdots + \; f_{nf}\, q^{-nf}$$

The model selection criteria used in this study are the polynomial degree, the final prediction error FPE, and the percentage of fit between the generated curve from measured data against that estimated by the model. However, from the statistical point of view there are other criteria that can be applied to evaluate and conclude whether a model is appropriate, such as CP (Mallows, 1964), CAT (Parzen, 1974), BIC (Sawa, 1978), and SBIC (Schwarz, 1978) [4].

2.4 Final prediction error – FPE

Known as final prediction Akaike error provides a measure of the quality of a model by simulating with a data set. This criterion indicates that the most accurate model is the one with the lowest FPE. The calculation of the FPE is given by

$$FPE = V \left(\frac{1 + {}^{d}/_{N}}{1 - {}^{d}/_{N}} \right) \tag{8}$$

where: *V* is the loss function.
D is the number of estimated parameters.
N is the number of values in the data set.

2.5 Percentage of fit

Percentage of fit represents the percentage of the output that the model can reproduce. It is expected to be a high setting of the fit percentage. This value is calculated with

$$BF = \left(1 - \frac{|y - \hat{y}|}{|y - \bar{y}|} \right) * 100 \tag{9}$$

where: y are the output validation data.
$\hat{y}$ are the output data estimated by the model.

2.6 Polynomial degree

This criterion seeks for a model whose structure is formed by least degree polynomials and presenting good characteristics of BF and FPE. Low order models are preferred in the analysis. An average between the polynomials of the structures taking into account both the degree of the polynomial and the number of polynomials of the structure is proposed in

$$GP = \frac{(x_1 + x_2 + x_3 + x_4 + x_5)}{5} \tag{10}$$

where: X1, X2, X3, and X4 represent the degrees of the polynomials used in the structures, and X5 the number of lags in the input. These parameters are represented in a matrix MIn_{x5}, being X_j the five components corresponding to the ith row of the matrix MI, with i = 1, 2, ... n (where n is the population size).

3 Multi-objective optimization

Most real-world problems require solving multiple objectives simultaneously [5]. In multi-objective optimization problems the following characteristics are found:

- The objectives may have different units of measure.
- There are conflicting objectives, which do not lead to a unique solution better than the set of solutions studied with regard to the expected objectives. Then a set of alternative solutions representing the best compromise between the objectives is obtained.
- A decision maker with information of preferences to choose the solution to implement is needed.

The methods of multi-objective optimization problems are classified into:

3.1 A priori methods

Decision making is done before searching solutions. Scaling processes to combine the different objectives into a single objective are often used. This scaling process is performed based on preference information that the decision maker suggests. This way a single-objective optimization problem is obtained.

3.2 A posteriori methods

A decision is made once the search ends. The optimization process is performed without considering preference information. This leads to an optimal Pareto set from which a solution satisfying the preferences for the problem is chosen.

3.3 Interactive or progressive methods

The decision making is realized during the search. This is accomplished by presenting a set of compromising solutions to a decision maker in each optimization step. Thus the decision maker can provide information to guide the search process in future iterations.

www.witpress.com, ISSN 1743-355X (on-line)

4 Multi-objective evolutionary algorithms

The first evolutionary algorithms (EAs) that consider multiple objectives were developed in the 90s. Multi-Objective Evolutionary Algorithms (MOEAs) have been established as a method for approximating the Pareto-optimal front [5]. Methods for working with multiple objectives using EAs can be classified into methods of first and second generation. In the first generation are both the initial algorithms that do not consider Pareto concepts and algorithms that considered Pareto concepts but do not have preservation mechanisms of good solutions. The second generation consists of Pareto-based EAs that incorporate forms of elitism. Algorithms as MOGA, NSGA, and NPGA belong to the first generation. In the second generation are found algorithms as SPEA and NSGA-II [5]. Evolutionary techniques have been implemented in the field of control process, where multi-objective optimization concepts are applied to a multi-criteria decision situation to develop computational methods for adjusting PID Controllers [6].

4.1 MAGO

The Multidynamics Algorithm for Global Optimization (MAGO) is a heuristics resulting from the combination of Lagrangian Evolution, Statistical Control, and Estimation of Distribution [7]. It does not use genetic operators but is based on statistics from the same evolving population. MAGO has only two parameters: number of generations and population size. Unlike other EAs, to get a larger exploration-exploitation balance and less likelihood to convergence to a local optimum, MAGO has three different dynamics for evolving the population. The cardinality of each dynamics changes in each generation, according to a rule inspired on methods of statistical control. The Emergent Dynamics is composed of improved elite which seeks solutions in a neighborhood near the best of all the individuals. This subgroup has the function of making faster convergence of the algorithm. The Crowd Dynamics is created by sampling from a uniform distribution, determined by the upper and lower limits of the second dispersion and the mean of the current population. This subgroup seeks possible solutions in a neighborhood close to the population mean. The Accidental Dynamics is the smaller one but has two basic functions: maintaining the diversity of the population, and ensuring numerical stability of the algorithm. Individuals of this subgroup are taken as samples from a uniform distribution throughout the searching space, similarly as in the initial population.

4.2 Application of evolutionary techniques

MAGO handles global evolution strategies, i.e. the use of genetic operators, such as crossover, is not raised because they required the analysis of each particular individual. MAGO through the three different dynamics produces new individuals in each generation. Each dynamics produces a subset of the new population. These three subgroups are the Emergent Dynamics made up of the Pareto front, Crowd Dynamics, and Accidental Dynamics. Figure 1 illustrates the way of determining the amount of individuals for each dynamics. The average of the current generation is a virtual individual calculated on purpose. The cardinality of the Emergent Dynamics corresponds to the number of elements within one standard deviation of the actual population. The cardinality of Crowd Dynamics is the difference between the first and second deviation. The number of remaining elements is the cardinality of the Accidental Dynamics.

Once the number of individuals within each dynamics is determined, MAGO proceeds to create individuals who will make up these groups and continues with the evaluation of new models. Figure 2 shows the flowchart for the evolutionary

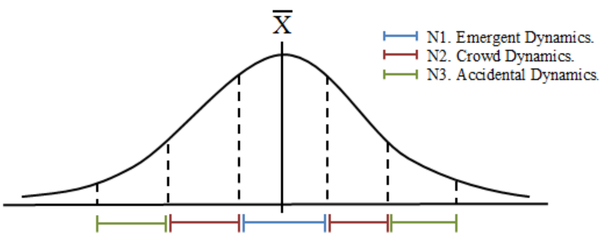

Figure 1: Cardinality of the dynamics for the generation of new populations.

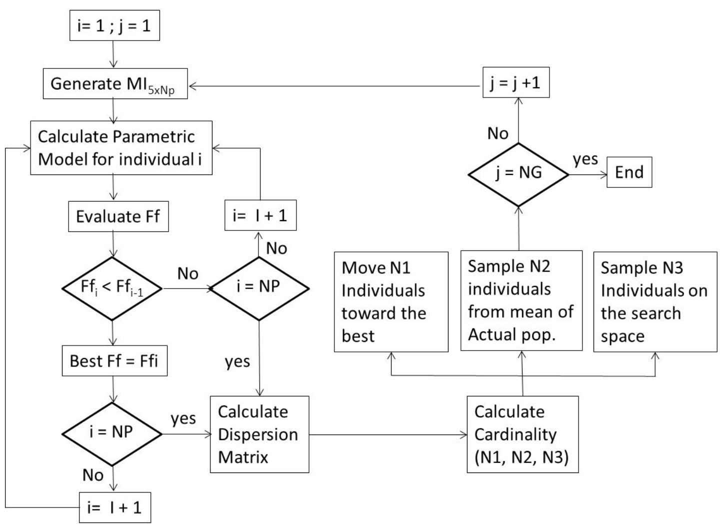

Figure 2: Flowchart for evolutionary multi-objective systems identification.

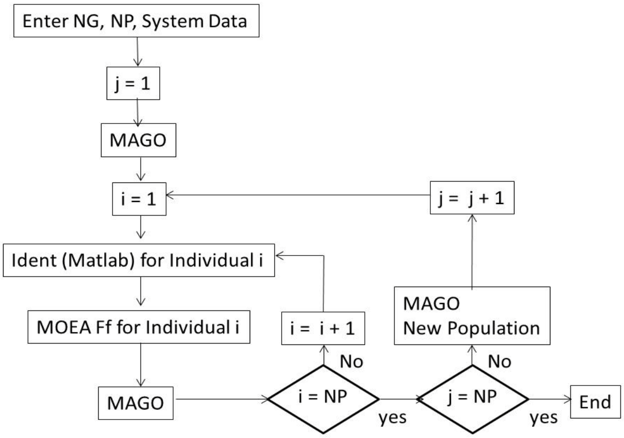

Figure 3: Software packages general flowchart.

multi-objective systems identification process. Figure 3 shows the flowchart of software packages. Following is the MAGO pseudo code:

1: j:= 0.
2: Random initial population with a uniform distribution over the search space.
3: **Repeat.**
4: Evaluate each individual with the fitness function.
5: Calculate the scattering matrix of the population.
6: Calculate cardinalities N1, N2, and N3 of the three dynamics G1, G2, and G3.
7: Select the N1 best individuals, move toward the best of all, make the displaced compete with their parents, and choose the best of them to the next generation j + 1.
8: Sample N2 individuals from a uniform distribution in the hyper rectangle [LB (j), UB (j)], and pass to the next generation j + 1.
9: Sample N3 individuals with a uniform distribution over the entire search space. Pass to the next generation j + 1.
10: j = j + 1.
11: **Until** to satisfy a stopping criterion.

5 Cases of study

The proposed method for identification systems was validated with both an academic device and an open standard data set. The academic device is the microclimate tool developed by National Instruments for academic purposes [8]. It consists of a bulb generating heat within an acrylic closed chamber, and also has a cooler that creates a constant air flow and a temperature sensor. The system is excited with a step of voltage of 3 volts. Through a data acquisition system a total of 120000 data with a sampling of 5 msec is obtained. From these, a total of 80000 data is used to find the best parametric model describing the dynamics of

www.witpress.com, ISSN 1743-355X (on-line)

the system using the algorithm MAGO. The other 40000 are used for validation purposes. Figure 4 shows the behavior of the system in time.

The data set of an installation that acts as a hair dryer is used as standard case of study [9]. Input data to the system is the voltage applied to the air heating device. Output data is the air temperature which is measured with a thermocouple. Figure 5 shows the hairdryer system behavior in time.

Multi-objective optimization and elitism techniques were used for creating parametric modeling of systems identification. Based on the criteria FPE, Fit Percent and Polynomial Degree a problem of multi-criteria decision is formulated. This problem deals with choosing the model presenting the greatest percentage of adjustment, the smallest FPE, and minimizing the degree of the polynomial. The steps of this algorithm are described below.

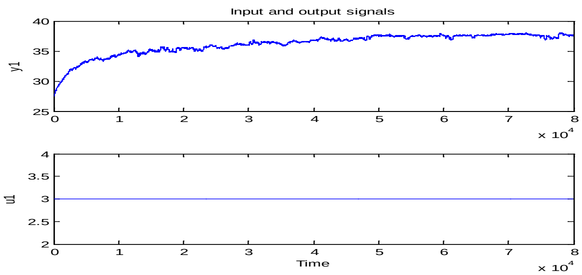

Figure 4: Microclimate system response (above) to step input of 3 volts (below).

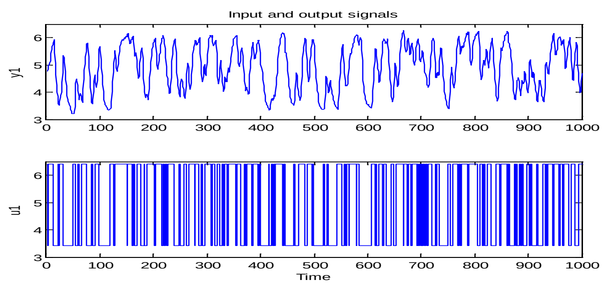

Figure 5: Hairdryer system response (above) to voltage steps at the input (below).

5.1 Creating the individual's matrix

The size of each population is 30 individuals. The studied models are the models OE, ARX, and ARMAX taking a total of 10 models on each structure. The parameters to be found are the FPE, the percentage of fit, and the polynomial degree of each individual or model. They represent, therefore, the individuals in the population within an array MI_{30x5}. Figure 6 shows the array of individuals from the first generation, being in each row the respective degrees of polynomial and individual delays.

Each row of the matrix MI represents an individual of the population. For example in the matrix MI there is the OE model [9 10 7] in the first row, the ARX [8 3 2] in the third row, and the ARMAX [5 2 9 2] in the last row. For the first generation rows from 1 to 10 represent OE models, rows from 11 to 20 represent ARX models, and rows from 21 to 30 represent ARMAX models. The number of individuals of each type of model changes in each generation but remains the total population of 30 individuals in each generation.

$$\boldsymbol{MI} = \begin{bmatrix} \mathbf{9} & \mathbf{10} & \mathbf{0} & \mathbf{0} & \mathbf{7} \\ \mathbf{2} & \mathbf{4} & \mathbf{0} & \mathbf{0} & \mathbf{11} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{8} & \mathbf{0} & \mathbf{3} & \mathbf{2} \\ \mathbf{0} & \mathbf{11} & \mathbf{0} & \mathbf{7} & \mathbf{3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{1} & \mathbf{7} & \mathbf{4} & \mathbf{0} & \mathbf{1} \\ \mathbf{5} & \mathbf{2} & \mathbf{9} & \mathbf{0} & \mathbf{2} \end{bmatrix}$$

Figure 6: Matrix of individuals.

After obtaining a population, a model evaluation proceeds. The criteria FPE, Percentage of Adjustment, and polynomial degree are used to select the most suitable models or individuals to represent the system.

5.2 Obtaining individual characteristics and coding of features

The toolbox Ident of Matlab is used to build the models. After deriving the models, features used in the EA (FPE, polynomial degree, and percentage of adjustment) are extracted. The encoding task is enabled since all values obtained with these characteristics are numerical.

For the degree of the polynomial, PD, an average among the polynomial degree of the structures present in the models considering both the number of polynomials and the degree of them is used, see equation (11).

$$\mathrm{PD} = \frac{\mathrm{Na+Nb+Nc+Nf+Nk}}{5} \tag{11}$$

where: Na, Nb, Nc, and Nf are the degrees of the polynomials of the model to evaluate, and Nk are the lags of the input.

Equation (12) shows how the percentage of fit adapts to be used in the program:

$$PA = \frac{pa}{10} \tag{12}$$

where: pa is the percentage of fit of the model.

A specific encoding for the numeric value of the final prediction error FPE is not applied, but its value is used directly.

5.3 Evaluation of characteristics with the fitness function

It is considered that a multi-objective optimization problem is mathematically solved when a Pareto optimal set has been found. There are three kinds of methods to solve a MOP (multi-objective optimization problem), a priori methods, a posteriori methods, and interactive methods or progressive [3]. The fitness function in equation (13) performs a scaling between the criteria PD and PA presented in equations (11) and (12), respectively, and the FPE.

$$Ff = (10 - PA) + PD + FPE \tag{13}$$

Most of the multi-objective optimization problems require as a solution a model offering the highest PA, also using the simplest structure and minimizing the FPE. Equation (13) presents an easy alternative taking advantage of a priori method reducing the system identification problem to a single-objective optimization problem. The selection model seeks the equation (13) minimization.

5.4 Obtaining the Pareto front

Individuals in the population with the lowest value Ff according to the fitness equation (13) are selected. These individuals form the Pareto front and are the most suitable models representing the system under study in each generation. The model with better characteristics is saved for the next generation, which is formed partially by a normal distribution around the best one. Figure 7 shows the Pareto front for the first and last generations for the microclimate system.

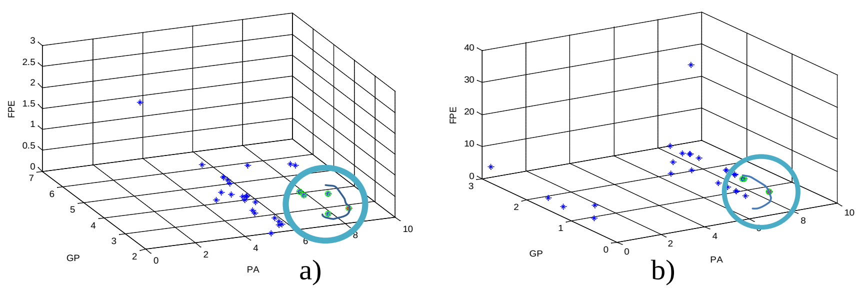

Figure 7: Pareto front for the microclimate system, the circle indicates the Pareto front estimated by MAGO. a) First generation, b) last generation.

6 Results

For the evaluation of the case study a reference analysis was constructed. This reference serves to compare the result with MAGO. The reference analysis involves a human expert evaluation about what is the most appropriate model to experimental data. This analysis was applied only to the micro system. For the standard data set there is not offered a model of reference.

According to the expert the best model describing the system is OE [2 2 0] with a PA of 85.67% and an FPE of 0.0707. Running the MAGO with 30 generations, also occurs as a result of the structure OE [2 2 0]. Additional to this, the algorithm also records another model as a feasible solution, the structure OE [1 2 1] with a PA of 85.66% and an FPE of 0.0707. From the comprehensive evaluation of the fitness function this model turns out to be more similar than the former. Although the computational effort is low considering the amount of analyzed models (around 1500 in 30 generations with a population of 30 individuals), this effort can be reduced with a proper adjustment of the fitness function. Adapting appropriately the acceptable limits of the polynomials degrees, it is possible to reach a solution with an evolution in fewer generations. Table 1 shows the best model selected in each generation. From generation 20, the OE [2 2 0] model remains as the dominant model in all populations until generation 30.

Table 1: Best individual on each generation for the microclimate system.

Generation	Best model	PD	PA	FPE	F_f
1…2	OE [5 6 1]	2.4	86.34%	0.0820	4.2977
3	ARX [1 2 1]	0.8	73.89%	0.1539	3.5649
4	OE [3 3 4]	2	86.34%	0.0985	3.4644
5	ARX [2 2 1]	1	78.55%	0.0128	3.2735
6	OE [3 5 1]	1.8	86.34%	0.0846	3.2508
7…11	OE [4 3 2]	1.8	86.34%	0.0789	3.2447
12…13	OE [4 3 1]	1.6	86.34%	0.1265	3.0924
14…15	OE [3 3 1]	1.4	86.34%	0.0746	2.8406
16	OE [1 4 1]	1.2	86.29%	0.0614	2.6322
17	OE [1 2 1]	0.8	85.66%	0.0707	2.3047
18	OE [1 2 1]	0.8	85.66%	0.0707	2.3047
19	OE [2 2 0]	0.8	85.67%	0.0707	2.3037
20…30	OE [2 2 0]	0.8	85.67%	0.0707	2.3037

For the HairDryer system, MAGO works with populations of 60 individuals and 30 generations. Results for the hairdryer are presented in Table 2. Best solution was obtained as OE model [3 3 0] with a PA of 87.96% and an FPE of 0.0101.

7 Conclusions

With the application of concepts of multi-objective optimization and multi-criteria decision an evolutionary computation procedure that can create and select parametric models for the identification of dynamical systems from a set of measured data has been developed. The algorithm MAGO successfully found the best model representing the microclimate system, model to which the expert had also arrived. Besides the evolutionary process proposed another equivalent model also meeting the selection criteria set for the problem under study.

Table 2: Best individual on each generation for the HairDryer system.

Generation	Best model	PD	PA	FPE	Ff
1	ARX [2 3 1]	1.2	67.51%	0.0546	2.6893
2⋯6	OE [5 1 0]	1.2	85.33%	0.0150	2.5764
7⋯10	OE [2 2 2]	1.2	87.57%	0.0170	2.4538
11	OE [3 2 1]	1.2	87.62%	0.0106	2.4489
12⋯13	OE [3 2 1]	1.2	87.62%	0.0106	2.4489
14	OE [2 3 1]	1.2	87.81%	0.0103	2.4293
15	OE [3 3 0]	1.2	87.96%	0.0101	2.4293
16⋯30	OE [3 3 0]	1.2	87.96%	0.0101	2.4142

Because the algorithm must apply statistical methods to calculate models in each generation this may have a high computational effort when large samples are taken into account. As future work the design of a module to conduct a pre-processing of information which decides the structures that are most appropriate for the system under study is considered. With these candidates structures the initial population could be randomly generated in order to reduce the processing time.

References

[1] Hernández J & Ospina J. *A multi dynamic algorithm for global optimization*, ELSEVIER, Mathematical and Computer Modelling 52. pp 1271-1278, 2010.

[2] Ljung L. *System identification Toolbox for use with Matlab*, (chapter 1-2). MathWorks.Inc. pp. 2-22, 1988.

[3] Isermann Rolf & Munchhof Marco. *Identification of Dynamic Systems*, ISBN 978-3-540-78878-2, Springer-Verlag Berlin Heidelberg, 2011.

[4] García C. *Stability of model selection criteria (in Spanish)*, Qüestuó 20(2). pp. 148-149, 1996.

[5] Von Lücken C. *Evolutionary Algorithms for multi-objective optimization: a comparative study in an asynchronous parallel environment (in Spanish).* Universidad Nacional de Asunción. pp 30-68, 2003.

[6] Reynosa G, Blasco X & Sanchis J. *Evolutionary multi-objective optimization and multi-criteria selection for control engineering (in Spanish),* Proceedings X Simposio CEA Ingeniería de Control UPC Barcelona. pp 1, 2012.

[7] Hernández J, Dorado J, Gestal M & Belén A. *Advances in Evolutionary Algorithms (in Spanish)* Inteligencia Artificial y Computación Avanzada. ISBN 84-934497-0-9, Fundac. Alfredo Brañas. Santiago de Compostela, Spain, 2007.

[8] National Instruments. QNET HVAC Trainer. Online (2013/03/14). http://www.ni.com/pdf/products/us/qnetdatasheet.pdf.

[9] Mechanical Systems. *Data of a laboratory setup acting like a hair dryer*. Departement Electrotechniek ESAT/SISTA, Online (2013/03/14). http://homes.esat.kuleuven.be/~smc/daisy/daisydata.html.

www.witpress.com, ISSN 1743-355X (on-line)