24th COBEM - 2017

24th ABCM International Congress of Mechanical Engineering
December 3-8, 2017, Curitiba, PR, Brazil

COBEM-2017-0542
META-HEURISTICS APPLIED TO SYSTEM IDENTIFICATION

Fábio Meneghetti Ugulino de Araújo
Alcemy Gabriel Vitor Severino
UFRN – Universidade Federal do Rio Grande do Norte
CT-DCA, Laboratório de Controle de Processos
Avenida Senador Salgado Filho, 3000
Campus Universitário – Lagoa Nova
59078-970 – Natal, RN – Brasil
alcemygvseverino@gmail.com, meneghet@dca.ufrn.br

Abstract. *There are different techniques for the problem of identifying systems, many are based on the least squares method and error reduction rate. In this paper meta-heuristics based system identification method are studied to automatically construct NARX models of nonlinear systems of unknown structure from observations of inputs and outputs. The meta-heuristics investigated to system identification are: Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Bat Algorithm (BA). A case study is identified to demonstrate the effectiveness of the meta-heuristics to compare it with traditional identification techniques.*

Keywords: *Meta-heuristic, System Identification, Selection of structures, NARX.*

1. INTRODUCTION

There are two method to represent a physical phenomenon based on mathematical models: the model constructed from physical equations that govern the phenomenon or observed data-based identification process.

Systems identification is an area that studies ways of modeling and analyzing systems from observed data (Aguirre, 2015). The model of a system is a mathematical equation used to answer questions about the system without the need to perform experiments (through a model we can calculate or decide how the system behaves) (Coelho and Coelho, 2015).

Models are classified in two types linear or nonlinear. In industry, most systems use approximate linear models of the processes they desire to control. The use of control techniques based on linear models is due in part to the simplicity of the models used to represent the behavior of the process. Linear systems do not exist in practice since all physical systems are nonlinear in some way (Golnaraghi and Kuo, 2012). In contrast, nonlinear control schemes, desire employ more realistic and therefore more complex models, replace the simplicity associated with linear techniques. Nonlinear models allow a more accurate "image" of the process when it is necessary.

There is a wide variety of nonlinear representations used in the systems identification, for example, Volterra series, the Hammerstein models and Wiener models. A practical difficulty in applying these representations to systems is that the number of parameters to be determined is very large (Aguirre, 2015).

In general, more concise models are desirable. For this reason, the NARX model (Nonlinear AutoRegressive with eXogenous inputs) was developed, which can represent a wide class of nonlinear systems with a small number of parameters (Chen and Billings, 1989; Leontaritis and Billings, 1985; Li and Jeon, 1993).

NARX models are linear in parameters and are functions of past inputs and outputs.

$$y(k) = F^l[y(k-1), y(k-2), \dots, y(k-n_y), \dots, u(k-d), \dots u(t-d-n_u)] + e(k) \tag{1}$$

where l is the degree of nonlinearity, u is the input, y is the output, d is the time delay, n_u and n_y are the orders in the input and output respectively, and $e(k)$ is the prediction error.

The number of terms can be determined by the following expression (Korenberg *et al.*, 1988):

$$\begin{aligned} n_\theta &= M + 1 \\ M &= \textstyle\sum_{i=1}^{l} n_i \\ n_i &= \frac{n_{i-1}(n_y + n_u + i - 1)}{i}, n_0 = 1 \end{aligned} \tag{2}$$

where n_θ is the number of terms in the model. In fact, n_θ increases sufficiently with the increase of the degree of nonlinearity l and of the orders n_y and n_u, becoming impracticable in some cases.

Fu and Li (2013) explain that there are several methods for the structure selection of mathematical models and that are divided into classical methods, based mainly on the algorithms of least squares method, correlation function and error reduction rate, and modern methods that use techniques of artificial neural networks, fuzzy logic, genetic algorithms, among others. For example, Li and Jeon (1993) applied the genetic algorithm in the selection of NARX model structure for systems identification.

In this paper meta-heuristics based system identification method are studied to automatically construct NARX models of nonlinear systems of unknown structure from observations of inputs and outputs. The meta-heuristics investigated to system identification are the Genetic Algorithm (GA), the Particle Swarm Optimization (PSO), and the Bat Algorithm (BA). A practical example is identified to demonstrate the effectiveness of the meta-heuristics to compare it with traditional identification techniques.

2. META-HEURISTICS APPLIED TO SYSTEM IDENTIFICATION

Meta-heuristics use a combination of random choices and the history of past results to conduct searches in the neighborhood of the search space, which reduces the possibility to stop in local minimums (Severino *et al.*, 2016).

We can understand the application of meta-heuristics in the structures selection of NARX models as a representation problem. In general, meta-heuristics use decimal and binary numbers, however, it is possible to apply different representations, such as string, sets of rules. In the case studied, the binary and decimal bases are more adequate.

According to the values of the delays n_y and n_u and the degree of nonlinearity l is generated the set of possible regressors. For example, for $n_y = n_u = 4$ and $l = 2$, a total of 44 candidate terms are obtained and 16383 structures of model possible. Thus, in this case, each candidate solution will be represented by a binary vector with 44 elements. The inclusion or exclusion of a term will be defined by the presence of the number 0 and 1 in each position of a given candidate solution of the meta-heuristic. If the term does not belong to the structure of the model will be associated the number 0 to the position corresponding to that term in the solution found by the meta-heuristic, and if it belongs, the number 1 will be associated. Thus, a binary vector with length equal to the number of candidate terms will inform which regressors will compose the model structure.

The meta-heuristic is responsible for generating and recalculating a set of binary vectors, in other words, the possible structures of the model to be identified. Once the method has been executed, it will select the most suitable model for the applied system.

The decimal base can also be used, for this, it is necessary the decimal-binary and binary-decimal conversion. Instead of binary vectors, integer numbers belonging to the range $[1, 2^{n_\theta} - 1]$ are used.

For example, for $n_y = n_u = 2$ and $l = 2$, we obtain the set of regressors with 14 terms. If the meta-heuristic indicates as the solution $\mathbf{i} = 2093$, this means that the terms $\{u^2(k-1),\ y(k-1)y(k-2),\ u(k-2),\ u(k-1),\ y(k-1)\}$ would compose the structure of the model, because $2093_{10} \rightarrow 00100000101101_2$ as shown in the Figure 1.

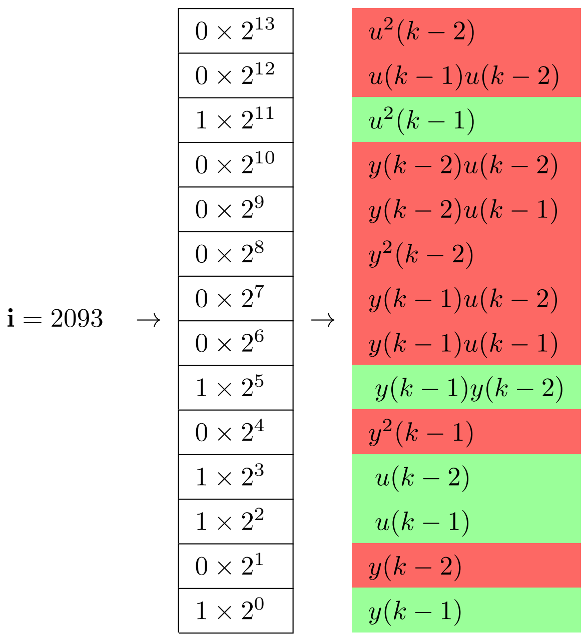


Figure 1: Set of candidate terms.

During the selection of the structure by the meta-heuristic, each member of the set of possible solutions (or possible structures) has its parameters estimated by the method of least squares, see Aguirre (2015, chap. 5). The parameters are then validated and each member is evaluated by the evaluation function. In this work, as can be seen in Equation 3, the evaluation function used in the meta-heuristic was Akaike's information criterion (Akaike, 1974):

$$f_a = AIC(n_\theta) = N \ln[\sigma^2_{\text{erro}}(n_\theta)] + 2n_\theta \tag{3}$$

where N is the length value of the collected data, σ^2_{erro} is the residue variance and n_θ is the number of parameters in the model.

Thus, from the binary coding system to represent the structure of NARX models, we can generalize the application of the meta-heuristics by the following steps:

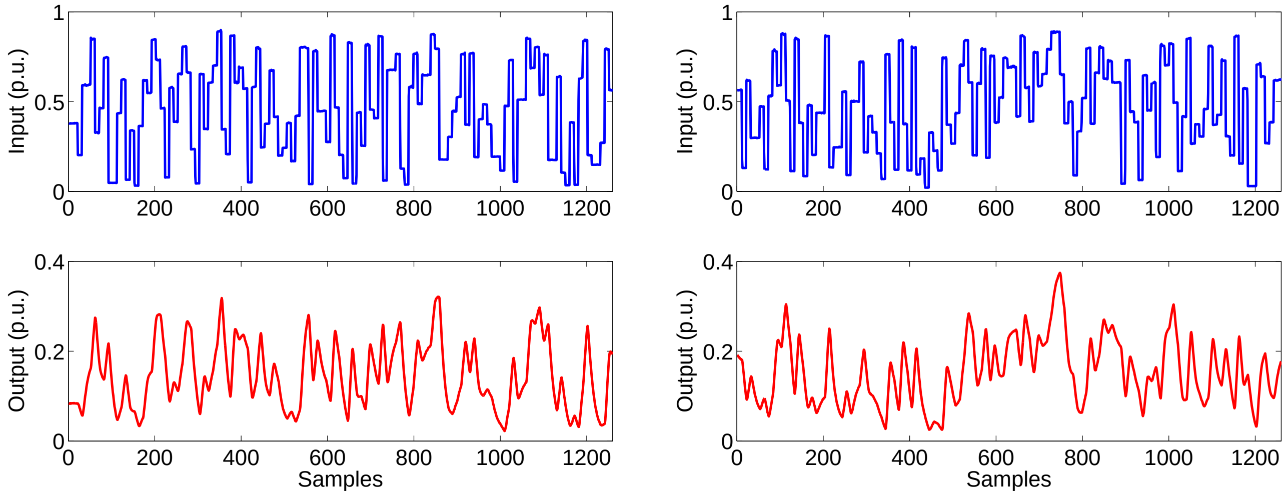
1. Define n_y, n_u and l, generate the candidate terms and read the data collected from the system.
2. Inform the number of binary vectors, number of executions and each parameter of the meta-heuristic.
3. Generate the initial set of solutions.
4. Evaluate the set of solutions.
5. If the maximum number of executions is reached, finalize the algorithm and inform the best model found. Otherwise, go to the next step.
6. Recalculate the set of solutions using the search strategies of the meta-heuristic, return to step 4.

The meta-heuristics GA, PSO and BA implemented in this work follow the classical algorithms Holland (1975), Eberhart *et al.* (1995) and Yang (2010), as described in Severino *et al.* (2016).

3. ELETRIC HEATER

The data used were collected of a thermoelectric system in which the input is the applied electrical voltage (p.u.) and the output is the measured electrical voltage at the terminals of a thermocouple (p.u.). The input signal is of the PRS type with a mean of 0.5 p.u., in which 1 p.u. corresponding to 136 V at the input voltage and at 998.51°C at the output temperature (Cassini, 1999).

The data was divided into two distinct sets, see Figure 2, the first one will be used for parameter estimation, while the second will be used for model validation. (Source: Cassini (1999) - data set <din3>).



(a) Estimation (b) Validation

Figure 2: Data used in identification.

Cassini (1999) obtained the following model NARX (Equation) for $n_y = n_u = 2$ and $l = 3$.

$$\begin{aligned} y(k) = \; & 1.3307y(k-1) - 0.3975y(k-2) - 0.1771y^2(k-1) + 0.0896y(k-1)u(k-1) \\ & -0.600y(k-1)u(k-2) - 0.0250y(k-2)u(k-1) + 0.0315u^2(k-1) \\ & +0.0120u(k-1)u(k-2) - 0.3481y^3(k-1) - 1.043y^2(k-1)u(k-1) \\ & +0.5037y^2(k-1)u(k-2) + 0.8896y(k-1)y(k-2)u(-1) + 0.4793y^3(k-2) \\ & -0.2768y^2(k-2)u(k-2) + 0.0046u^3(k-1) \end{aligned} \tag{4}$$

The linear terms of noise proposed by Cassini (1999) were used to minimize the polarization of the estimated parameters so that for simulation criteria they can be omitted. The use of $l > 3$, $n_y > 2$ and $n_u > 2$ may be considered unnecessary for the system identified herein because of its low degree of nonlinearity.

4. RESULTS

Each execution of the meta-heuristics contained a set of 30 possible solutions, calculated over 50 iterations. The meta-heuristics were applied the same set of possible initial solutions, where each member contains only one of the candidate terms: the first member receives the first term, the second member receives the second term, etc. If the number of terms was greater than the number of possible solutions, then the process is repeated from the initial term.

The values of the parameters of the metaheuristics (Table 1) were defined through a series of tests and observations until their values were considered adequate, in other words, the parameters were configured empirically.

Table 1: Parameters for each meta-heuristic

Meta-heuristic	Parameter	
GA	t_{mut}	0.5%
PSO	ω	0.10
	ϕ_1	1.40
	ϕ_2	1.40
BA	f_{min}	0.00
	f_{max}	0.04
	α	0.20
	γ	0.20

The initial velocities of the PSO and BA were defined as 0, it means the particles and bats are initially stopped; in addition, the rate r_0 of pulse emission and the loudness A_0 of the BA had their values set at 0.1 and 2, respectively.

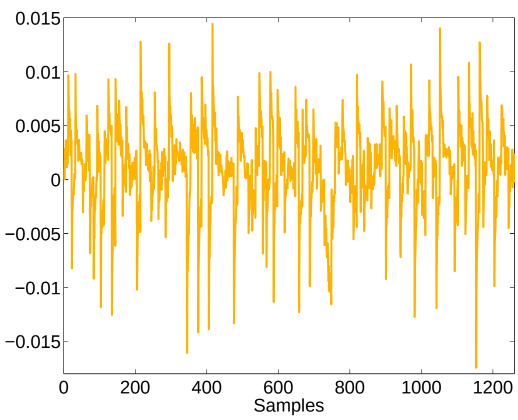
The NARX models, for $n_y = 2$, $n_u = 2$ and $l = 3$, obtained by GA, PSO and BA are given by Equations 5, 6 and 7, respectively.

$$\begin{aligned} y(k) = \; & 1.4228y(k-1) - 0.4600y(k-2) - 0.0049y(k-1)u(k-2) + 0.0144u^2(k-1) \\ & +0.0059u^2(k-2) + 0.0012u(k-1)u^2(k-2) \end{aligned} \tag{5}$$

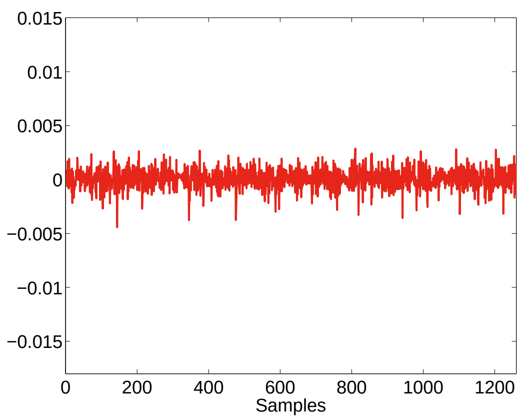
$$\begin{aligned} y(k) = \; & 1.3987y(k-1) - 0.4396y(k-2) + 0.0015u(k-2) + 0.0293y(k-1)u(k-1) \\ & -0.0282y(k-2)u(k-2) + 0.0103u^2(k-1) + 0.0087u^2(k-2) \end{aligned} \tag{6}$$

$$\begin{aligned} y(k) = \; & 1.3755y(k-1) - 0.4165y(k-2) + 0.0020u(k-2) - 0.0134y(k-1)y(k-2) \\ & +0.0078y(k-1)u(k-1) + 0.0823y(k-1)u(k-2) - 0.0810y(k-2)u(k-2) \\ & +0.0099u^2(k-1) + 0.0071u(k-1)u(k-2) \end{aligned} \tag{7}$$

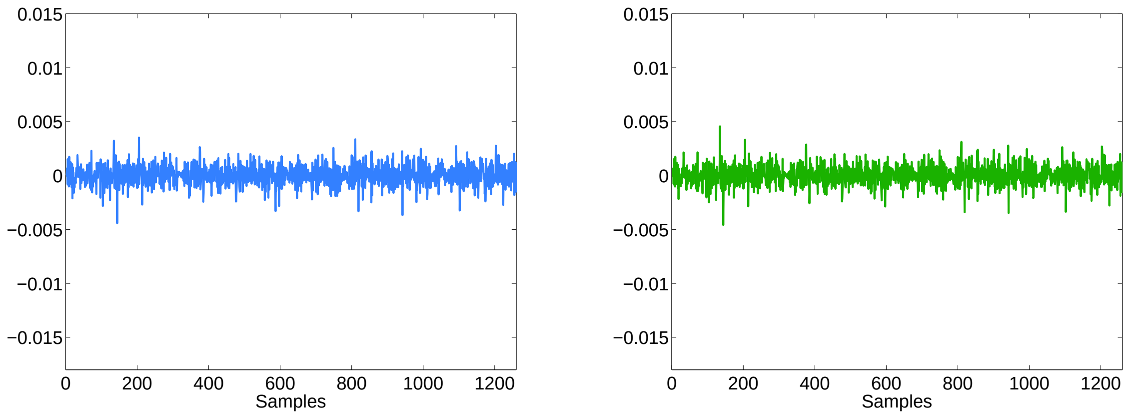
The error, $e(k) = y(k) - \hat{y}(k)$, obtained by each of the models is shown in the Figure 3



(a) Cassini.



(b) GA.



(c) PSO. (d) BA.

Figure 3: Difference between actual and estimated output.

To make possible a better distinction between the errors obtained by the GA, PSO and BA models, Figure 4 details the curves contained in Figure 3, the error presented by the Cassini (1999) model has a larger amplitude, therefore the same was not included.

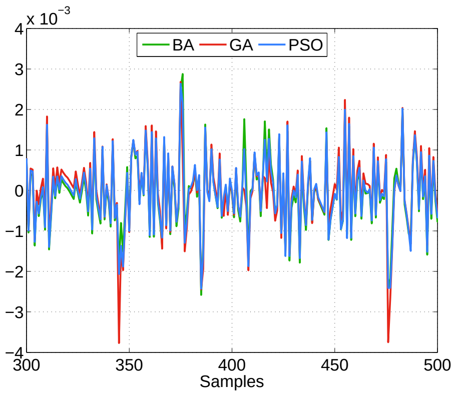


Figure 4: The error curves of the models obtained by the meta-heuristics in detail.

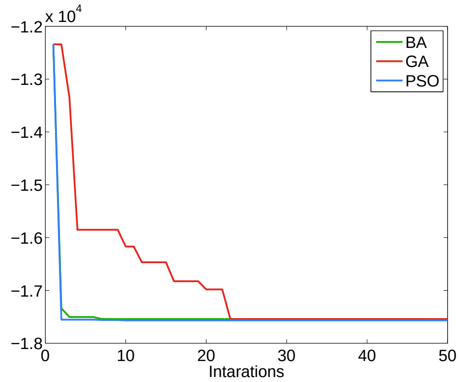


Figure 5: The evolution of the AIC value for the best solution found by the meta-heuristic at each iteration.

As a comparison criterion, the sum of the squares of the difference between the observed and estimated values, given by the equation below, was calculated.

$$J_e = \sum_{k=1}^{N} [y(k) - \hat{y}(k)]^2 \quad (8)$$

where N is equal to the number of samples.

Each execution registers the number of model structures that have been visited by meta-heuristics, called n_s. The maximum value of n_s is equal to $2^{n_r} - 1$, where n_r is the number of possible regressors. We can calculate the percentage of the search space that was traversed by the meta-heuristic, called $n_{s_{\%}}$, based n_s.

Table 2 contains the AIC, J_e, n_θ, n_s and $n_{s_{\%}}$ values for each model evaluated.

Table 2: Evaluation indices.

Model	AIC	J_e	n_θ	n_s	$n_{s_{\%}}$
Cassini	-1.396336×10^4	2.050185×10^{-2}	15	-	-
GA	-1.754273×10^4	1.123215×10^{-3}	6	686	$3.99 \times 10^{-8}\%$
PSO	-1.756827×10^4	1.096015×10^{-3}	7	945	$5.50 \times 10^{-8}\%$
BA	-1.755865×10^4	1.100920×10^{-3}	9	1287	$7.49 \times 10^{-8}\%$

According to Figures 3 and 4, and Table 2, we can affirm that the models found by metaheuristics have a lower level of error. The model estimated by the PSO is the one with the best performance, however, the other two meta-heuristics obtained AIC and J_e values very close to the PSO.

Observing the number of regressors of the selected models, the estimated by GA was the model with the lowest number of terms, 6 regressors in total, a number almost 3 times smaller than the model proposed by Cassini (1999). Both the PSO and BA found models with 7 and 9 regressors, respectively. With regard to the linearity of the terms, the GA found a model with 3 linear terms and 2 nonlinear terms (three quadratic and one cubic). The PSO defined a model with 3 linear terms and 4 nonlinear (all quadratic) terms. The BA selected a model with 3 linear and 6 nonlinear terms (all quadratic).

With respect to the evolution of the AIC value (Figure 5), the PSO and BA show an abrupt decrease compared to the GA. In the first iterations, the PSO and BA already found models with AIC values very close to the final models presented, behavior that occurred more slowly in GA. However, this does not mean that the previous models had the same structure as the final models.

5. CONCLUSION

In this work, the meta-heuristics known as genetic algorithm, particle swarm optimization, and bat algorithm were applied in the problem of the structures selection for NARX model in the identification of a nonlinear system. The performance of these meta-heuristics was compared with the model proposed by Cassini (1999) for the modeling of an electric heater. The error presented by the models found by these techniques are smaller than the model obtained by Cassini (1999). The best model analyzed was found by the PSO, however, the models selected by the GA and the BA presented results very close to the PSO. With regard to the number of terms, the best result was that of GA, which obtained a model with only 6 regressors, but among the analyzed models it was the one with the highest nonlinearity ($l = 3$). Thus, the results demonstrate that meta-heuristics are efficient strategies in solving the problem of the selection of structures for NARX models. In future works, other meta-heuristics can be considered with the objective of making better comparisons between the traditional techniques applied in the selection of model structures. As well as analyzing the use of different evaluation functions for the studied problem.

6. ACKNOWLEDGEMENTS

CAPES, CNPq e PPGEMECA/UFRN for financial support and grant of scholarships.

7. REFERENCES

Aguirre, L.A., 2015. *Introdução à identificação de sistemas–Técnicas lineares e não-lineares aplicadas a sistemas reais*. Editora UFMG, Belo Horizonte, MG.

Akaike, H., 1974. "A new look at the statistical model identification". *IEEE transactions on automatic control*, Vol. 19, No. 6, pp. 716–723.

Cassini, C.C.S., 1999. *Estimação recursiva de características estáticas não lineares utilizando modelos polinomiais NARMAX*. Master's thesis, Universidade Federal de Minas Gerais, Belo Horizonte.

Chen, S. and Billings, S.A., 1989. "Representations of non-linear systems: the NARMAX model". *Int. J. Control*, Vol. 49, No. 3, pp. 1013–1032.

Coelho, A.A.R. and Coelho, L.S., 2015. *Identificação de sistemas dinâmicos lineares*. Editora UFSC, Florianópolis, SC.

Eberhart, R.C., Kennedy, J. *et al.*, 1995. "A new optimizer using particle swarm theory". In *Proceedings of the sixth international symposium on micro machine and human science*. New York, NY, Vol. 1, pp. 39–43.

Fu, L. and Li, P., 2013. "The research survey of system identification method". In *Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2013 5th International Conference on*. IEEE, Vol. 2, pp. 397–401.

Golnaraghi, F. and Kuo, B.C., 2012. *Sistemas de Controle Automático*. LTC, Rio de Janeiro, 9th edition.

Holland, J.H., 1975. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence.* MIT Press, Cambridge, MA, 1st edition.

Korenberg, M., Billings, S., Liu, Y. and McIlroy, P., 1988. "Orthogonal parameter estimation algorithm for non-linear stochastic systems". *International Journal of Control*, Vol. 48, No. 1, pp. 193–210.

Leontaritis, I. and Billings, S.A., 1985. "Input-output parametric models for non-linear systems part i: deterministic non-linear systems". *International journal of control*, Vol. 41, No. 2, pp. 303–328.

Li, C.J. and Jeon, Y., 1993. "Genetic algorithm in identifying non linear auto regressive with exogenous input models for non linear systems". In *American Control Conference, 1993*. IEEE, pp. 2305–2309.

Severino, A.G.V., Araújo, Í.B.Q., Linhares, L.L.S. and Araújo, F.M.U., 2016. "Metaheurísticas para estimação de parâmetros na identificação de sistema não lineares utilizando modelos NARMAX polinomiais". *IX Congresso Nacional de Engenharia Mecânica*.

Yang, X.S., 2010. "A new metaheuristic bat-inspired algorithm". In *Nature inspired cooperative strategies for optimization (NICSO 2010)*, Springer, pp. 65–74.

8. RESPONSIBILITY NOTICE