# The Performance of Artificial Bee Colony (ABC) in Structure Selection of Polynomial NARX Models

Azlee Zabidia, Nooritawati Md Tahira and Ihsan Mohd Yassin
*Faculty of Electrical Engineering,*
*University Technology Mara,*
*40450 Shah Alam*
*salam.uitm.edu.my*

***Abstract*—System Identification (SI) is a discipline of building a mathematical model of dynamic systems based on its input and output data. The process of SI is generally divided into structure selection, parameter estimation and model validation. This paper attempts to address the structure selection issue in SI, where the objective is to select the most representative set of regressors to represent the system. However, the selection process must obey the principle of parsimony, where the structure must be as small as possible, yet has the ability to represent the system well. We propose a binarized modification of the Artificial Bee Colony (ABC) algorithm to perform structure selection of a Nonlinear Auto-Regressive with eXogenous (NARX) model on a Direct Current (DC) motor. We compare this implementation with the Binary Particle Swarm Optimization (BPSO) algorithm in terms of solution quality and convergence consistency. The results indicate that the ABC algorithm excelled in terms of convergence consistency with similar solution quality to BPSO algorithm.**

***Index Terms*—System Identification (SI); Artificial Bee Colony (ABC); Binary Particle Swarm Optimization (BPSO); Nonlinear Auto-Regressive with eXogenous (NARX).**

## I. INTRODUCTION

SI is a process of building a mathematical model of a dynamic system from observed input/output data [1, 2]. It has been widely use to model observations in various fields such as manufacturing, engineering, aviation, astronomy, ecology and economics [1, 2]. This field has been receiving significant attention because an increasing number of complex observations and discoveries [1, 3]that require progressive approach to represent beyond traditional modeling techniques.

The process of SI is generally divided into structure selection, parameter estimation and model validation. This paper attempts to address the structure selection issue in SI, where the objective is to select the most representative set of data to represent the system. However, the selection process must obey the principle of parsimony, where the structure must be as small as possible, yet has the ability to represent the system well.

We propose a binarized modification of the Artificial Bee Colony (ABC) algorithm to perform structure selection of a Nonlinear Auto-Regressive with eXogenous (NARX) model on a Direct Current (DC) motor dataset. The ABC algorithm is an optimization algorithm that mimics the intelligent behavior of bee colonies in finding the best food source around it perimeter, by dividing the responsibilities of the swarm and working together to achieve a common goal [4]. The performance of ABC is compared with another swarm-inspired optimization algorithm, the Binary Particle Swarm Optimization Algorithm (BPSO) [5], with the focus of solution quality and convergence consistency. This work is motivated by previous works by [4, 6, 7] that suggest that the ABC algorithm outperforms PSO in terms of local and global optimization.

## II. LITERATURE REVIEW

### *A. NARX*

NARX is a SI model which represents the output behavior of a system based on its past inputs and outputs [8]. The NARX model can be constructed using various methods, such as polynomials[8], Multilayer Perceptrons (MLP) [9], and Wavelet ANNs (WNN) [10], although the polynomial approach is the only method that can explicitly define the relationship between the input/output data.

The polynomial representation of the NARX model for a given input–output series is [8]:

$$y(t) = \sum_{m-1}^{n_p} P_m \theta_m + \varepsilon(t) \tag{1}$$

where $n_p$ is the number of terms in the polynomial expansion, $P_m$ is the $m$-th regression term with $P_1 = 1$, and $\theta_m$ is the $m$-th regression parameter. $P_m$is formed by a combination of input, output and residual terms. In matrix form, identification involves the formulation and solution of the LS problem:

$$P\theta + \varepsilon = y \tag{2}$$

where $P$ is a $n \times m$ regressor matrix, $\theta$ is a $m \times 1$ coefficient vector, and $y$ is the $n \times 1$ vector of actual observations. $P$is arranged such that its columns represent the $m$ lagged regressors. $\varepsilon$is the white noise residuals.

### *B. The ABC algorithm*

The ABC algorithm is based on the cooperative behavior of natural bees in the swarm. In a bee swarm, there are three types of bees that have specific roles. Scout bees evaluate the fitness of the solution (termed nectar amount), and this information is shared with onlooker bees waiting in the hive. After the initial search, all scout bees now become employed bees. The employed bees go to the food sources (solutions) in its memory and determines the neighboring food sources to evaluate the nectar amount. If the neighboring food source contains a better solution, the new position is kept. Otherwise the old position is maintained [4, 11].

The information of the new or existing nectar amount is then relayed to awaiting onlookers when the employed bees return. Onlooker bees then select a food source depending on nectar amount relayed. If the nectar amount increases (solution approaching objective), the probability which that food source is selected is higher. The employed bees which carrying high nectar amount will attract onlooker bees toward it food sources position.

After selecting a potential food source from employed bees, the onlooker bee goes toward the direction and evaluate the neighboring food source. Similar to employed bees, if the neighboring food source contains a better solution, the new position is kept [4, 11]. Otherwise, the old position is maintained. The process is repeated between employed and onlooker bees until the food source is finished. Once this happens, scout bees are again sent out to discover new food sources. In ABC, the activation of scout bees is controlled by how many iterations in which no better-quality food sources are discovered [11].

In order to binarize the ABC algorithm, we follow the concept outlined by [5] by representing the bee positions as “probabilities of change” rather than the actual solution. This generates a binary sequence that is used to select or reject regressor terms, thus performing the structure selection process on the model.

## III. Methodology

All experiments were performed on a personal computer with 3.10GHz Intel Xeon E3-1220 v3 microprocessor and 4GB of Random Access Memory (RAM). The operating system used was Linux Mint XFCE version 17.1 with MATLAB 2014a as the development platform. The flowchart for the structure selection process is shown in Figure 1 and the parameter settings for BABC and BPSO is shown in Table 1. These parameters were selected to test the robustness of both algorithms under different initialization and exploration conditions.

The dataset used was a Direct Current (DC) motor [12, 13]. The dataset is a Single Input Single Output (SISO) relating a nonlinear relationship between the input voltage and the angular velocity of the motor. The model order of the system is 2 as stated in [12, 13].

The dataset was preprocessed using four configurations prior to the experiment:

1. No magnitude scaling, 50:50 training and testing division ratio using block division method (PP1).
2. No magnitude scaling,50:50 training and testing division ratio using interleaving division method (PP2).
3. Magnitude scaling between -1 and 1, 50:50 training and testing division ratio using block division method (PP3).
4. Magnitude scaling between -1 and 1, 50:50 training and testing division ratio using interleaving division method (PP3).

The regressor matrix was created based on the model order of two. A total of 14 regressor terms were generated,namely [u(t-1), u(t-2), y(t-1), y(t-2), u(t-1)*u(t-1), u(t-1)*u(t-2), u(t-1)*y(t-1), u(t-1)*y(t-2), u(t-2)*u(t-2), u(t-2)*y(t-1), u(t-2)*u(t-2)]. The number of possible combinations was$2^{14}$ = 16,384.

After the regressor matrix was created, the BABC and BPSO algorithms were used to select the best possible structure guided by the Akaike Information Criterion (AIC), Final Prediction Error (FPE) and Model Descriptor Length (MDL) [14] as the fitness function.

Several tests, namely the One Step Ahead (OSA) prediction, residual plot, correlation tests and residual histogram analysis, were performed to validate the model.

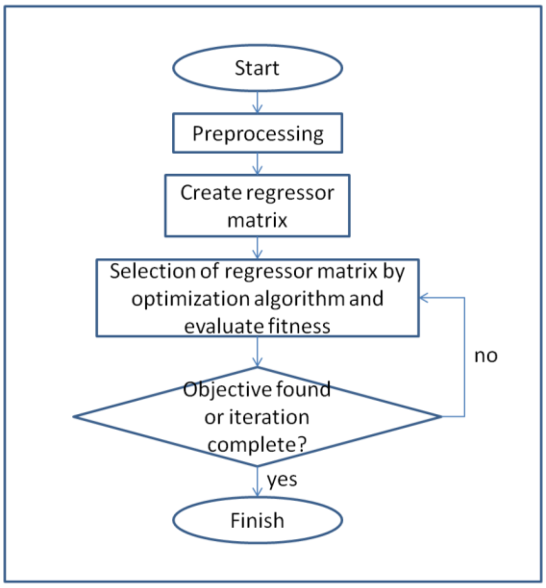


Figure 1: The BABC and BPSO optimization process

## IV. Result and Discussion

The optimal structure selection results were obtained using FPE as the fitness criterion and PP2 as the preprocessing method. Both BABC and BPSO obtained similar solutions with FPE of $2.13\times10^{-7}$. The training and testing MSE values from the optimal solutions were $3.39\times10^{-7}$ and $3.29\times10^{-7}$, respectively. A summary of results from all fitness criteria is shown in Table 2.

The OSA prediction for BABC and BPSO training and testing sets are shown in

Figure 2 and Figure 3, respectively.

The OSA prediction results show a close fit between the predicted results and original data with high r-squared value. These observations indicate that the model was able to approximate the dynamics of the original system well. This observation is also confirmed based on the small magnitude of the residuals, as shown in

Figure 4.

However, for the model to be accepted, the residuals of the model need to exhibit properties similar to white noise. This can be validated using correlation and histogram tests, shown in

Figure 5 to

Figure 6.A majority of the correlation coefficients reside between the 95% confidence limit, while only a minority of coefficients exceeding the confidence limit by a very small margin. Additionally, the histogram test results in Figure 7 follow a bell-shaped curve, indicating a Gaussian distribution. Both observations indicate that the residuals are random and uncorrelated, thus exhibiting white noise properties. Based on this, the model is considered as valid and

acceptable.

The convergence consistency of both BABC and BPSO were analyzed by examining the final FPE fitness values based on different parameter combinations. A total of 45 and 180 parameter combinations were analyzed for BPSO and BABC, respectively. BABC had more parameter combinations because of an additional limit parameter, which is not applicable for BPSO.

The quality of solution is similar as both algorithms converged to the optimal solution. However, the consistency of BABC in finding the optimal solution was 98.8%, while BPSO managed to find the optimal solution 80% of the time (Figure 8). This observation shows that BABC was far superior to BPSO in terms of convergence consistency.

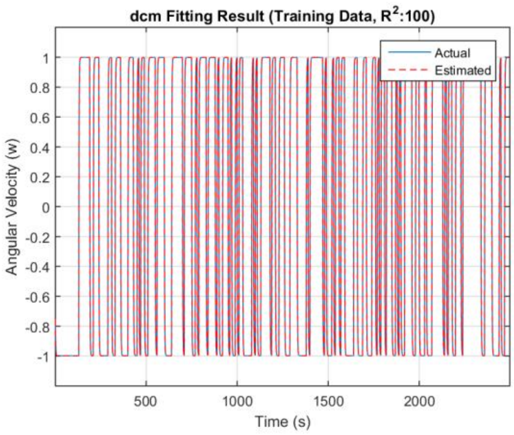


Figure 2: OSA prediction (training set)

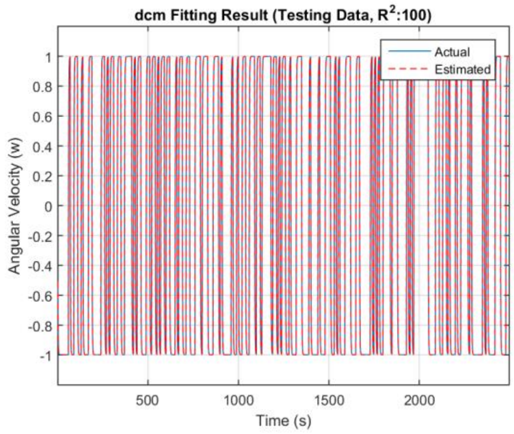


Figure 3: OSA prediction (testing set)

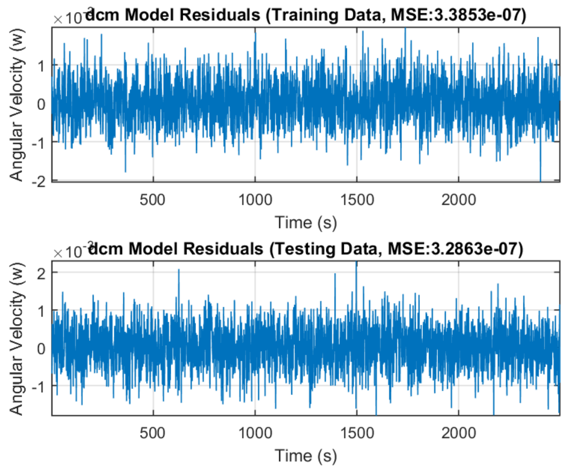


Figure 4: Residual plot of prediction model

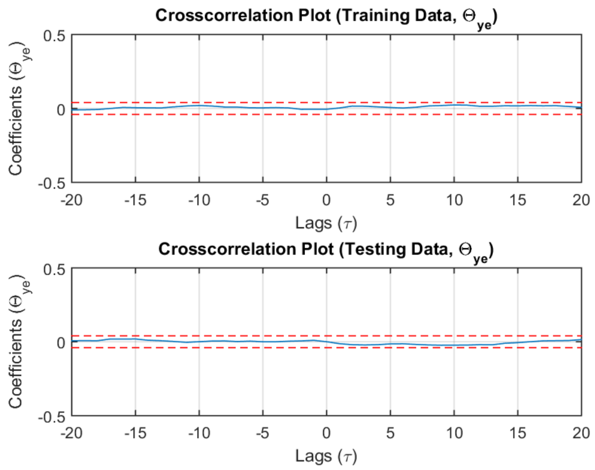


Figure 5: Cross-correlation plot of residuals and model output

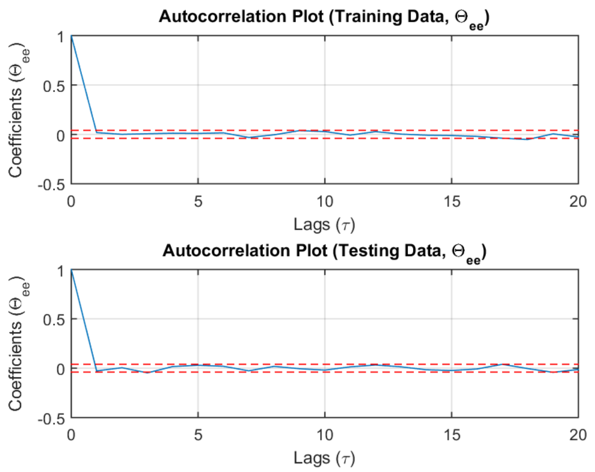


Figure 6: Auto-correlation plot of residuals

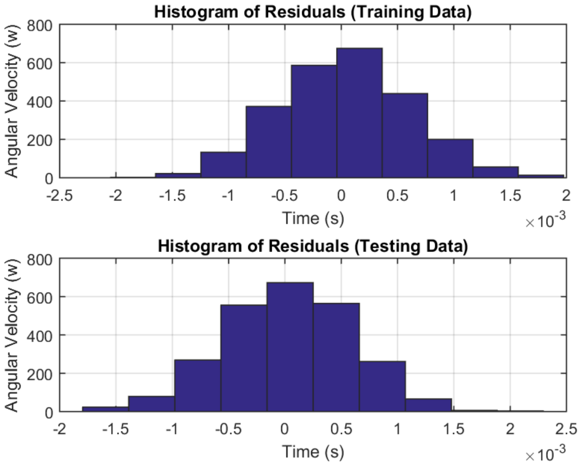


Figure 7: Histogram of residuals

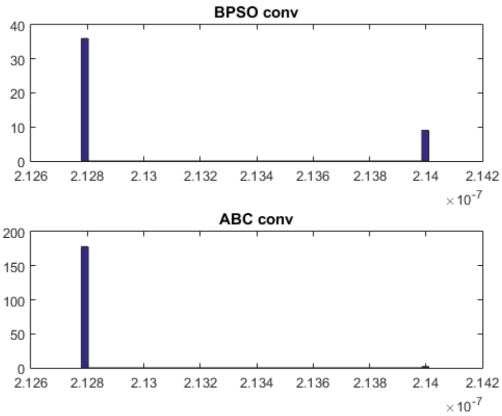


Figure 8: Convergence consistency of BPSO and BABC

## V. CONCLUSION

The optimal solution achieved for both BPSO and BABC was $2.13\times10^{-7}$ by using FPE as fitness criterion with PP2 preprocessing method. Both algorithms were able to find the optimal result from 16,384 potential solutions. The optimal solution had passed all the necessary tests for it to be considered a valid model. However, in terms of consistency, it appears that the BABC algorithm can locate the solution more consistently relative to BPSO, thus making it a superior algorithm for structure selection.

## ACKNOWLEDGMENT

The authors would like to express our gratitude to Universiti Teknologi Mara (UiTM) for funding through its Lestari Grant (600-RMI/DANA 5/3/LESTARI (5/2015)), the Ministry of Higher Education Malaysia for its support through the MyPhD scholarship program, and the Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, for its equipment and infrastructure support.

## REFERENCES

[1] Li F. and PengfeiL., "The Research Survey of System Identification Method," *in 5th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, pp. 397-401, 2013.

[2] LingT. G., RahmatM. F., and HusainA. R., "System identification and control of an Electro-Hydraulic Actuator system," *in IEEE 8th International Colloquium on Signal Processing and its Applications (CSPA)*, pp. 85-88, 2012.

[3] L. Guangjun, X. Xiaoping, and W. Feng, "Identification of a kind of nonlinear system," *in 2011 Seventh International Conference on Natural Computation (ICNC)*, pp. 1730-1733, 2011.

[4] KarabogaD. and BasturkB., "On the performance of artificial bee colony (ABC) algorithm," *Applied Soft Computing*, vol. 8, pp. 687-697, 2008.

[5] Kennedy J. and EberhartR. C., "A discrete binary version of the particle swarm algorithm," *in Proc. 1997 Conf. Systems, Man, Cybernetics, Piscataway, NJ*, pp. 4104–4108, 1997.

[6] TuranogluE., ÖzceylanE., and KiranM. S., "Particle swarm optimization and artificial bee colony approaches to optimize of single input-output fuzzy membership functions," *Proceedings of the 41st International Conference on Computers & Industrial Engineering*, pp. 542-547.

[7] AkayB., "A study on particle swarm optimization and artificial bee colony algorithms for multilevel thresholding," *Applied Soft Computing. 13*, pp. 3066-3091, 2013.

[8] AmisigoB. A., GiesenN. v. d., RogersC., AndahW. E. I., and FriesenJ., 2007.Monthly streamflow prediction in the Volta Basin of West Africa: A SISO NARMAX polynomial modelling," Physics and Chemistry of the Earth. 33(1-2):141-150.

[9] RahimN. A., 2004.The Design Of A Non-Linear Autoregressive Moving Average With Exegenous Input (NARMAX) For A DC Motor. MSc, Faculty of Electrical Engineering, UniversitiTeknologi MARA, Shah Alam.

[10] Billings S. A. and WeiH.-L., 2005.A New Class of Wavelet Networks for Nonlinear System Identification. IEEE Trans. Neural Networks. 16(4): 862-874.

[11] DixitG. P., DubeyH. M., PanditM., and PanigrahiB. K., 2011. Artificial bee colony optimization for combined economic load and emission dispatch.in International Conference on Sustainable Energy and Intelligent Systems (SEISCON 2011). 340-345.

[12] I. M. Yassin, M. N. Taib, N. A. Rahim, M. K. M. Salleh, and H. Z. Abidin, 2010. Particle Swarm Optimization for NARX structure selection - Application on DC motor model.in IEEE Symposium on Industrial Electronics & Applications (ISIEA). 456-462.

[13] YassinI. M., TaibM. N., Abdul AzizM. Z., RahimN. A., TahirN. M., and JohariA., 2011. Identification of DC motor drive system model using Radial Basis Function (RBF) Neural Network.in IEEE Industrial Electronics and Applications (ISIEA). 13-18.

[14] Mendez E. M. A. M. and BillingsS. A., 2001.An alternative solution to the model structure selection problem. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on. 31: 597-608.

[15] EspinozaM., SuykensJ. A. K., and MoorB. D., 2005.Kernel Based Partially Linear Models and Nonlinear Identification. IEEE Trans. Automatic Control. 50(10): 1602-1606.