International Review of Automatic Control (I.RE.A.CO.), Vol. 7, N. 2
ISSN 1974-6059 March 2014

Evolutionary Computation in System Identification: Review and Recommendations

Md Fahmi Abd Samad

Abstract – *Two of the steps in system identification are model structure selection and parameter estimation. In model structure selection, several model structures are evaluated and selected. Because the evaluation of all possible model structures during selection and estimation of the parameters requires a lot of time, a rigorous method in which these tasks can be simplified is usually preferred. This paper reviews cumulatively some of the methods that have been tried since the past 40 years. Among the methods, evolutionary computation is known to be the most recent one and hereby being reviewed in more detail, including what advantages the method contains and how it is specifically implemented. At the end of the paper, some recommendations are provided on how evolutionary computation can be utilized in a more effective way. In short, these are by modifying the search strategy and simplifying the procedure based on problem a priori knowledge.*

Keywords*: System Identification, Model Structure Selection, Evolutionary Computation, Genetic Algorithm, Review*

Nomenclature

d. c.	Constant level
M*	Parsimonious model
M	General model
n_u	Maximum orders of lag for input
n_y	Maximum orders of lag for output
p_c	Crossover probability
p_m	Mutation probability
u	Input variable
y	Output variable
$\mu + \lambda$	Parent + children selection strategy

I. Introduction

System identification (SI) is a method of recognizing the characteristics of a system, thus producing a quantitative input-output relationship that explains or resembles the system's dynamics. The procedure involves the interpretation of observed or measured data into physical relationship, often and easily interpreted in the form of mathematical models [1]. The field has long been an interest to many and according to [2], profound results in linear SI have been around since 1970s.

SI is also related to modelling. There are two types of modelling and these are *a priori* modelling and *a posteriori* modelling [1].

An *a priori* modelling stresses the construction of a mathematical model of a system based on the knowledge of the scientific interdependence of variables whether mechanical, electrical, chemical or others. An *a posteriori* modelling, on the other hand, does not require any theoretical knowledge of the system.

The construction of a model in an *a posteriori* modelling is also referred to as black-box modelling. In black-box modelling, the variables enter the model independently via a prediction procedure. It is crucial to have a method that is able to perform this task with an optimized prioritization on aspects such as accuracy of model, simplicity of model, cost, time consumption, ease and also flexibility in method implementation.

This paper attempts to recollect various methods that had been tried in performing model structure selection.

The paper provides a systematic introduction to SI and model structure selection in Section 2. It points out, in particular, certain important considerations to clarify the requirements of the selection process.

Then, the methods used to perform them are reviewed in Section 3 with regards to certain significant feature of the method. In Section 4, evolutionary computation (EC) is looked into in more detail. Certain characteristics of EC are emphasized towards helping readers in making choices of their usage. In Section 5, some recommendations are provided on how EC can be applied more effectively. Section 6 concludes the content of the paper.

II. Model Structure Selection of System Identification

The procedure of identification can be divided into several distinctive steps. These steps are data acquisition, model structure selection, parameter estimation and model validity tests. The flow chart is given as Fig. 1.

Throughout the years, SI has evolved largely in terms

Manuscript received and revised February 2014, accepted March 2014

of procedures and methods in the steps above, mainly in model representation and parameter estimation stages.

Examples of practical nonlinear SI can be found in [3] and [4].

II.1. Steps of Model Structure Selection

Model structure selection stage can be divided into two sequential steps:

(i) Selection on the type of model to represent the system.

(ii) Construction of the correct or optimum model structure.

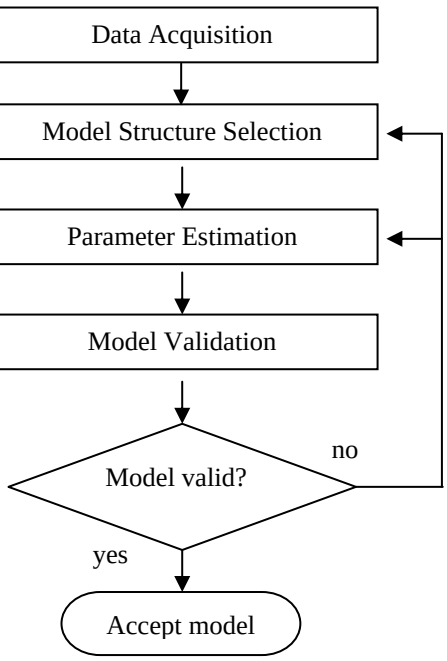

Fig. 1. Typical system identification flow

In today's literature, various types of models are proposed for system modelling. The most common classification of models is based on whether the model represents time-invariant or time-varying systems. For time-invariant systems, difference equation models are usually preferred. For time-varying systems, among popular choices are cascaded block model, neural network, wavelet network and cellular automata.

Difference equation models can be divided into linear and nonlinear models. For a general representation of various classes of linear difference equation model, an ARMAX (AutoRegressive Moving Average with eXogenous input) model is used.

It includes the representation of AutoRegressive model, Moving Average model and all other combinations with or without an exogenous variable.

Nonlinear models give much richer possibilities in describing systems and have better flexibility when inferring from a finite data set. Nonlinear models are used when higher degree variables are used to describe a system.

Examples of nonlinear difference equation model include NARX (Nonlinear AutoRegressive with eXogenous input), NARMAX (Nonlinear ARMAX) and NOE (Nonlinear Output Error) models. Their polynomials can be identified by their linear counterparts with addition of higher degree terms [5].

The NARMAX model is a general representation for finite nonlinear difference systems with obvious advantages over functional series expansions [6]. It is also proven to provide a general representation of various classes of nonlinear systems including bilinear, output-affine and rational models [7]. Another important feature of the NARMAX model is that it is linear-in-the-parameters allowing easy implementation of parameter estimation.

Linear and nonlinear functions can also be present simultaneously in a cascaded form. The most easily distinguished nonlinear model representation in the class of cascaded block models are the Wiener and Hammerstein models [1], [5]. A popular type of nonlinear model today is the neural network.

Originally, the neural network is developed with the metaphor of brain synapses. There are many types of neural network such as feedforward, back-propagation and recurrent network [8], [9]. A popular class of neural network is radial basis function network [10],[11].

By utilizing wavelet frame decomposition in its activation function, the model is known as wavelet network. (see e.g. [12]. Not to be missed out in the group of nonlinear model is the cellular automata.

Investigations include detection of the neighbourhood structure and extraction of rules using forward orthogonal least squares algorithm [13] and using genetic algorithm [14]-[16].

II.2. Considerations and Requirements

When choosing a model type, certain considerations need to be accounted. Most literature stresses on the purpose of modelling [1], [5]. The purposes may vary from establishing a simple regulating model to understanding a system's complete dynamics. In [17], the characteristics of an optimum model include reflection of all physical knowledge and external behaviours regarding the system at-hand towards providing an acceptable level of approximation. If possible, it should also lead to a simple selection and estimation algorithm.

[18] listed the considerations as flexibility, parsimony, algorithm complexity and properties of the criterion function.

Once the type of model is specified, a finite boundary of specification for the model needs to be determined.

For a discrete-time model, this refers to the process of determining the maximum lags of input, output, noise and time delay from the information of input and output sequences [19].

Some strategies on how model specification determination can be accomplished are given in [5]. With this specification selection, a model set, M, is defined

where, in this set, the models are made up of all candidate variables and terms. The next step is to determine the variables and terms that significantly explain the system's dynamics. Another widely discussed priority is parsimony of model structure [18].

In the principle of parsimony, a model need not only be adequate in representing a system, but also exhibits simple interrelationship among the variables and terms to ease the process of control design

Therefore, a parsimonious model, M*, that is formed from an original model, M, exhibits less complexity and has the property such that M* is a subset of M, written as $M^* \subset M$. Note that a model set defined here only considers the presence and absence of variables and terms where terms are used to refer to a variable of different lags and/or nonlinearity. Further differentiation of selected models becomes available when parameter estimation step is implemented [5].

III. Model Structure Selection Methods

Several specific methods applied for model structure selection are tabulated in Table I. However, these methods contain weaknesses which among all include its reliance on additional information either user-defined or heuristically-defined.

TABLE I
METHODS APPLIED TO MODEL STRUCTURE SELECTION

Method	Example Publications/ References	Remarks
Regression	[20]	Multiple calculation and recomputation and also undesirable for many variables and terms
Orthogonal Least Squares (OLS) based on Error Reduction Ratio (ERR)	[13], [21]	ERR depends on order of inclusion of terms and terms may be redundant
OLS-ERR and all subset search	[22]	Heavy computation when possible terms are many
Forward OLS-ERR	[12], [23]	Requires stepwise order inclusion procedure and ERR cut-off value
Hypothesis testing	[24]	Reliance on non-conflicting significance levels
Cross-validation	[25]	Requires separation proportion of data and multiple computation
Genetic Algorithm (minor modifications)	[26]-[28]	Premature convergence, inability for fine local tuning and inadequacy for handling model parsimony
Genetic Algorithm based on migration and artificial selection	[29], [30]	Slow convergence or requires tuning of algorithm parameters, fixed number of variables or terms and inability for fine local tuning
Memetic Algorithm with Penalty Function	[31]	Reliance on convergence of initial algorithm
Modified Genetic Algorithm	[32], [33]	Requires tuning of algorithm parameters including grouping proportion
Deterministic Mutation Algorithm	[34]	Premature convergence

Further downfalls of the regression methods are discussed in [22]. The orthogonal least squares (OLS) has an advantage over the regression methods by allowing independent parameter estimation.

However, it relies on optimum decomposition of the regressor matrix and the selection of threshold values of ERR to distinguish significance. The main disadvantage is that the change in the values of ERR for an identified model is usually gradual causing difficulty in justifying the level of significance and thus deciding the number of variables and terms to be included in the final model.

Furthermore, the models identified using OLS can become unstable when excessive iteration is made [22].

Among the methods in Table I, the last five (beginning with genetic algorithm) falls under the category of EC. The use of developed genetic algorithm has been shown to be better than the OLS method [31], [32].

IV. Evolutionary Computation

The EC is defined as a study of computational systems which uses ideas and gets inspirations from natural evolution and adaptation [35]. It is widely known as a group of stochastic search and optimization methods those operates on the metaphor of natural biological evolution as introduced by the famous Charles Darwin in 1859. It also stresses on species evolution according to the principle of survival-of-the-fittest [36]-[38].

Biological simulation has been used for optimization since the 1950s while the emergence of EC methods date to the 1960s and 1970s when parallel computation begin to receive significant development [39]. There are four distinctive methods in EC those are:

(i) Genetic algorithm (GA).
(ii) Evolution strategies (ES).
(iii) Evolutionary programming (EP).
(iv) Genetic programming (GP).

The first three methods are known as evolutionary algorithms while all four methods (GA, ES, EP and GP) altogether is termed EC since 1991 [40]. The name evolutionary algorithms (EA) and EC are sometimes used interchangeably although EA is more frequently used in referring to algorithmic procedures.

Origins and characteristics of these methods can be referred from [41]-[43] for genetic algorithm and [44], [45] for genetic programming. References such as [36], [46], [47] are advised for readings on EC as a whole and for other EC methods.

IV.1. Applications of Evolutionary Computation in Modelling

As a group of search and optimization method, EC has been applied in various fields. These are reported and characterized in numerous literatures [43], [48], [49].

For example, [48] characterized the possible applications of EC into five categories namely planning, design, simulation (and identification), control and lastly,

classification. Reviews on field-specific applications are in the field of electric power systems [50], electromagnetics [51], manufacturing [52], production and operations management [53], control systems engineering [54], traffic control [55], reservoir operation [56], robotics, artificial life, biology and medicine [57].

[58] provides a review of genetic algorithm as a control tool.

The applications of EC in modelling can be divided into model structure selection, parameter estimation and concurrent implementation of both applications. The model structure selection stage requires a robust method that is able to search, within the search space, the global optimum model structure. Although its usage in the area is still quite restricted, some examples of its usage in model structure selection stage are available in literature [10], [29]-[32], [59]. The first four references applied modifications/refinement of GA to difference equations model. Ref [59] considered block-oriented models while [10] applied GA to radial basis function network configuration. EC has also been applied for cellular automata (CA) neighbourhood term selection [15]-[17].

An example of the usage of EC for multivariable system modelling is present in [60].

In parameter estimation application, where the model structure is already defined, EC is used to search for the optimum parameter values in system behaviour modelling.

The models used in this application include difference equations model [61], [62], state-space model [63], fuzzy rules model [64], pole-zero transfer function model [65] and block-oriented models [66]. [66] demonstrated the use of both GA and ES for different piecewise approximation parameters. It has also been used in parallel with neural network modelling [67], [68].

Other applications integrate both the selection of a model structure and the estimation of parameters concurrently. Basically, for such an application, the evolution of a potential optimum model is carried out by sequentially identifying the terms for the model and then, the parameters of the terms [27], [28], [69]. Whilst some algorithms wait for the structure to be selected before parameter estimation, others search for both the structure and parameters simultaneously [26], [70].

IV.2. Procedural Selection in Evolutionary Computation Application

In EC applications, various procedures and operators are suggested and used. Table II provides a list of choices of algorithm parameters among academia, specifically for SI purpose. This section reviews some of the procedures of EC those are identified as being notably argumentative lately.

IV.3. Representation and Population Initialization

The representation stage refers to the mapping of problem components into EC individual representation.

This can be implemented either by using the real values of solutions also called the phenotype or different representative values such as binary ones also called genotype [84], [85]. Although different genetic operators apply for different types of representations, the transformation between binary and real-valued vectors is fairly easy [86].

Besides binary coding, Gray coding, an alternative of {0,1} cardinality representation, has also been used [43], [87]. Once the representation is completed, the population is initialized. Most population initialization is carried out randomly and very few researchers used heuristic information since this depends on the problem domain [53]. Particularly for model structure selection, [29] and [30], who used binary numbers, recommended a representation method where a solution is represented by a fixed number of allowable substrings. Each substring represents a term based on a decoding scheme. A scheme for a second degree NARX model, for example, is based on the following sequence:

$$\begin{aligned} &\text{d. c.}, y(t-1),...,y(t-n_y),u(t-1),...,u(t-n_u),\\ &y^2(t-1),...,y^2(t-n_y),u^2(t-1),...,u^2(t-n_y),\\ &y(t-1)u(t-1),...,y(t-n_y)u(t-n_u) \end{aligned} \quad (1)$$

where d. c. represents a constant level while n_y and n_u represent the maximum orders of lag for output y and input u, respectively. However, in this method, the number of regressors for the model is fixed and thus disabling neither further accuracy achievement through more regressors' inclusion nor parsimony search. Another method is to assume each component of a representation to represent a variable or a term of a model where the allele or representation is either 1 for present or 0 for omission [26], [28], [31], [32].

IV.4. Constraint-Handling Techniques

Constraint-handling technique is important in differentiating feasible solution from unfeasible ones based on constraints specification. Common techniques that can be used for overcoming constraints are as follows [88]:

(i) Setting the representations to sensible values.
(ii) Encoding/manipulating the solution to prevent invalid solution.
(iii) Reducing/penalizing performance value through evaluation for not meeting constraints.

These techniques are connected to what known as repair algorithm, decoder and penalty function, respectively in [86]. A new approach is the multiobjective optimisation [54]. Among these techniques, the penalty approach is the simplest. It is achieved by penalizing less feasible solutions by comparison of constraint conformance, such as the compliance to the principle of parsimony and reduction of model complexity, through penalty functions.

TABLE II
CHOICES OF PROCEDURES IN EC FOR SYSTEM IDENTIFICATION

Publication	Problem Representation (Real or binary)	Selection Strategy	Crossover Type	Mutation Type	Remarks
[26]	Binary and real number	Roulette-wheel and elitist	One-point	Hierarchical	Mutation operator is problem-dependent
[28]	Binary and real number	Fitness-based grouping	One-point and specialized	Random bit and gene regeneration	-
[31]	Binary number	Fitness-based grouping	One-point	Bit-flipping	Include regeneration and specialized local search
[32]	Binary number	Fitness-based grouping	One-point	Bit-flipping	Include regeneration
[66]	Binary and real number	Roulette-wheel and $(\mu + \lambda)$	One-point	Bit-flipping and Gaussian	-
[67]	Binary number	Fitness-proportional	Sequential-search-based	Not specified	-
[68]	Binary number	Roulette-wheel	One-point	Bit-flipping	-
[69]	Real number	Ranking	Arithmetical	Non-uniform	Include regeneration and local search by adaptation of search heuristics
[70]	Binary number	Tournament	Multi-point	Bit-flipping	Include regeneration
[71]	Real number	Roulette wheel	Uniform	Random and incremental/ decremental	-
[72]	Real number	Tournament	One-point	Adaptive	Include variable elitism
[73]	Binary number	Roulette-wheel	Uniform	Understood as bit-flipping	-
[74]	Understood as real number	Tournament	One-point	Not specified	Hybrid
[75]	Real number	$(\mu + \lambda)$	Not specified	Deterministic	Mutation operator is analytical
[76]	Real number	$(\mu + \lambda)$	Global discrete	Gaussian	Include elitism and hybrid
[77]	Real number	Pareto ranking	Gaussian and adaptive	Adaptive	-
[78]	Real number	$(\mu + \lambda)$	Not specified	Deterministic	Mutation operator is analytical
[79]	Binary number	Roulette-wheel and elitist	One-point	Random bit	-
[80]	Binary number	Ranking, roulette-wheel and artificial (fitness-based grouping)	One-point and multi-point	Random bit, cyclic non-uniform and local non-uniform	Include random migration and regeneration
[81]	Real number	Tournament	Not specified	Not specified	-
[82]	Real number	$(\mu + \lambda)$	Multiple recombination	Gaussian	-
[83]	Real number	Tournament	One-point and two-point	Gaussian and random	-

The use of penalty functions are practical for many types of constraints [89] and also traced in information criterion for evaluation of possible models, such as in Akaike's Information Criterion (AIC), B-information criterion and ϕ-information criterion [19]. Examples of the use of penalty functions in GA are found for nonlinear programming [90], radial basis function network modelling [10] and difference equation modelling [31], [32].

V. Potential Areas for Improvement

The overall development of EC usage shows two distinct areas for further improvement. One of the areas is the usage of modified/specialized operators to increase its search efficiency in term of exploitation of good information and exploration of search space. Another area is by simplifying its usage based on problem-specific knowledge. These areas need to be looked into to overcome problems like premature convergence and high computation time.

V.1. Increment of Search Efficiency Via Modified/Specialized Operators

As seen in Table II, in order to overcome the problem of premature convergence in EC, researchers resort to local search techniques and specialized operators or procedures. Examples of local search techniques incorporated in EC are simulated annealing [63] and memetic algorithm [31].

Some researchers designed specialized crossover operator to achieve more efficient search [91]. More recently and frequently, the EC community relies on specialized mutation operator in their algorithms either as a standalone operator or for fine-tuning of search [26], [34], [75], [78], [92], [93].

Another way of increasing the search efficiency of EC is by incorporating operators that allow more diversified search. The implementation of such operators allows more points to be evaluated during the search in order to avoid convergent to local optima. Among strategies are individual or population regeneration [69] and grouping of individuals to different operators [28], [32], [80].

[80] refers to the groups as species and uses different types of crossovers and mutations for each group.

Overall, an efficient search is always characterized by a good balance between fine-tuning and population diversification.

V.2. Simplification of Procedure and Evaluation Via Problem-Specific Knowledge

One of the issues in EC implementation is large computation time due to complicated or non-optimum setting of procedures and parameters. Research is wide in search of the optimum setting of the parameters especially when constraints are present [30], [32], [94], [95]. The crossover probability p_c, for example, was selected to be 0.6 by [96] and in the interval [0.75, 0.95] as reported in [97].

Other results of crossover probability include recommendation in the interval [0.45, 0.95] depending on offline or online usage [94] and $p_c \in$ [0.05, 0.6] depending on the characteristics of model, like order of lags [32]. Results do not always support each other when a change to a single parameter is made or different optimization problems are attempted as suggested in [87] and [98]. Furthermore, its applicability to practical problems is questioned due to the reason that some of these findings were made based on simple function optimization problems [99].

The same applies to other parameters like mutation probability p_m and population size. Poor selection of these parameters may cause GA to either converge prematurely or too slowly.

More specialized operators may be desirable in reducing computation time by the achievement of quicker convergence but as problem-specific situation is encountered, the all-purpose EC methods should be simplified. By restating the requirements and understanding the constraints of the problem, more effective procedures and objective function can be developed.

VI. Conclusion

This paper provides a review on the application of EC in SI. Stages of SI are briefly explained and some common methodologies in implementing these stages are described. One of the stages is model structure selection which refers to the task of determining the variables and terms in a process model that adequately describes the dynamics of the process.

Besides providing accurate prediction, the model should also be parsimonious.

As one of the method for model structure selection, EC is explained. However, its application in model structure selection is restricted by the factor of algorithm parameter tuning that may contribute to unnecessarily high computation time, poor exploration and premature convergence.

Potential areas for improvement of EC method are identified as modification of search strategy and simplification of procedure. These areas are believed to be promising in proposing a more efficient way of obtaining an accurate and parsimonious model.

Acknowledgements

The author wishes to acknowledge the funding by Universiti Teknikal Malaysia Melaka through Skim Latihan Akademik Bumiputera and Grant PJP/2012/FKM(20C)/S01117 during the undertaking of this work.

References

[1] R. Johansson, *System Modeling & Identification* (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1993)

[2] X. Hong, R. J. Mitchell, S. Chen, C. J. Harris, K. Li and G. W. Irwin, Model Selection Approaches for Non-Linear System Identification: A Review. *International Journal of Systems Science. Vol. 39*, n. 10, pp. 925-946, 2008.

[3] S. A. Billings and M. B. Fadzil, The Practical Identification of Systems with Nonlinearities, *Proceeding of the 7th IFAC Symposium of Identification and System Parameter Estimation*, York, United Kingdom, July, 1985, pp. 155-160.

[4] M. Thomson, S. P. Schooling, and M. Soufian, The Practical Application of a Nonlinear Identification Methodology. *Control Engineering Practice. Vol. 4*, n. 3, pp. 295-306, 1996.

[5] L. Ljung, *System Identification: Theory for the User* (2nd ed., Upper Saddle River, New Jersey: Prentice Hall PTR, 1999).

[6] S. A. Billings, Identification of Nonlinear Systems. In S. A. Billings, J. O. Gray and D. H. Owens, (Eds.). *Nonlinear System Design* (London, United Kingdom: Peter Peregrinus Ltd., 1984, 30-45).

[7] S. Chen and S. A. Billings, Representations of Non-Linear Systems: The NARMAX Model. *International Journal of Control. Vol. 49*, n. 3, pp. 1013-1032, 1989.

[8] S. A. Billings, H. B. Jamaluddin and S. A. Chen, Comparison of the Backpropagation and Recursive Prediction Error Algorithms for Training Neural Networks. *Mechanical Systems and Signal Processing. Vol. 5*, n. 3, pp. 233-255, 1991.

[9] M. Negnevitsky, *Artificial Intelligence: A Guide to Intelligent System* (Harlow, England: Pearson Education Limited, 2002).

[10] S. A. Billings and G. L. Zheng, Radial Basis Function Network Configuration Using Genetic Algorithms. *Neural Networks. Vol. 8*, n. 6, pp. 877-890, 1995.

[11] R. Rashid, H. Jamaluddin and N. A. Saidina Amin, Application of Radial Basis Function Network in Modeling the Tapioca Starch Hydrolisis, *Proceedings of the Second International Conference on Artificial Intelligence in Engineering & Technology*, Kota Kinabalu, Sabah, Malaysia: Universiti Malaysia Sabah, Aug 3-5, 2004, pp. 897-902.

[12] S. A. Billings and H. L. Wei, A New Class of Wavelet Networks for Nonlinear System Identification. *IEEE Transactions on Neural Networks. Vol. 16*, n. 4, pp. 862-874. 2005.

[13] S. A. Billings and Y. Yang, Identification of the Neighborhood and CA Rules From Spatio-Temporal CA Patterns. *IEEE Transactions on Systems, Man and Cybernetics – Part B; Cybernetics. Vol. 33*, n. 2, pp. 332-339, 2003.

[14] S. A. Billings and Y. Yang, Identification of Probabilistic Cellular Automata. *IEEE Transactions on Systems, Man, and Cybernetics*

– Part B: Cybernetics. Vol. 33, n. 2, pp. 225-236, 2003.
[15] Y. Yang and S. A. Billings, Neighbourhood Detection and Rule Selection from Cellular Automata Patterns. *IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans. Vol. 30*, n. 6, pp. 840-847, 2000.
[16] Y. Yang and S. A. Billings, Extracting Boolean Rules from CA Patterns. *IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics. Vol. 30*, n. 4, pp. 573-581, 2000.
[17] G. C. Goodwin and R. L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis* (New York: Academic Press, Inc., 1977).
[18] T. Söderström and P. Stoica, *System Identification* (London: Prentice Hall International (UK) Ltd., 1989).
[19] S. M. Veres, *Structure Selection of Stochastic Dynamic Systems: The Information Criterion Approach* (New York: Gordon and Breach Science Publishers, 1991).
[20] N. R. Draper and H. Smith, *Applied Regression Analysis.* (3rd ed., New York: John Wiley and Sons, Inc., 1998).
[21] M. Korenberg, S. A. Billings, Y. P. Liu and P. J. McIlroy, Orthogonal Parameter Estimation Algorithm for Non-linear Stochastic Systems. *International Journal of Control. Vol. 48*, n. 1, pp. 193-210, 1988.
[22] E. M. A. M. Mendes and S. A. Billings, An Alternative Solution to the Model Structure Selection Problem. *IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans. Vol. 31*, n. 6, pp. 597-608, 2001.
[23] H. L. Wei, S. A. Billings and J. Liu, Term and Variable Selection for Non-Linear System Identification. *International Journal of Control. Vol. 77*, n. 1, pp. 86-110, 2004.
[24] I. J. Leontaritis and S. A. Billings, Model Selection and Validation Methods for Non-linear Systems. *International Journal of Control. Vol. 45*, n. 1, pp. 311-341, 1987.
[25] P. Stoica, P. Eykhoff, P. Janssen and T. Söderström, Model-Structure Selection by Cross-Validation. *International Journal of Control. Vol. 43*, n. 6, pp. 1841-1878, 1986.
[26] T. G. Yen, C. C. Kang and W. J. Wang, A Genetic Based Fuzzy-Neural Networks Design for System Identification, *IEEE International Conference on Systems, Man and Cybernetics, Vol. 1*, Big Island of Hawaii, Oct 10-12, 2005, pp. 672-678.
[27] Y. Mitsukura, M. Fukumi, N. Akamatsu and T. Yamamoto, A System Identification Method Using a Hybrid-Type Genetic Algorithm, *Proceedings of the 41st SICE Annual Conference (SICE 2002), Vol. 3*, Osaka, Japan, Aug 5-7, 2002, pp. 1598-1602.
[28] A. Sakaguchi and T. Yamamoto, A Study on System Identification Using GA and GMDH Network, *The 29th Annual Conference of the IEEE Industrial Electronics Society (IECON'03), Vol. 3*, Roanoke, Virginia, Nov 2-6, 2003, pp. 2387-2392.
[29] G. C. Luh and G. Rizzoni, Nonlinear System Identification Using Genetic Algorithms with Application to Feedforward Control Design, *Proceedings of the American Control Conference. Vol. 4*, Philadelphia, Pennsylvania, June 24-26, 1998, pp. 2371-2375.
[30] G. C. Luh and C. Y. Wu, Nonlinear System Identification Using Genetic Algorithms. *Proceedings of the Institution of Mechanical Engineers - Part I: Journal of Systems and Control Engineering. Vol. 213*, n. 2, pp. 105-118, 1999.
[31] R. Ahmad, H. Jamaluddin and M. A. Hussain, Selection of a Model Structure in System Identification Using Memetic Algorithm, *Proceedings of the Second International Conference on Artificial Intelligence in Engineering & Technology*, Kota Kinabalu, Sabah, Malaysia: Universiti Malaysia Sabah, Aug 3-5, 2004, pp. 714-720.
[32] R. Ahmad, H. Jamaluddin and M. A. Hussain, Model Structure Selection for a Discrete-Time Non-Linear System Using Genetic Algorithm. *Proceedings of the Institution of Mechanical Engineers - Part I: Journal of Systems and Control Engineering. Vol. 218*, n. 2, pp. 85-98. 2004.
[33] H. Jamaluddin, M. F. Abd Samad, R. Ahmad and M. S. Yaacob, Optimum Grouping in a Modified Genetic Algorithm for Discrete-Time, Non-Linear System Identification. *Proceeding of Institution of Mechanical Engineers Part I: Journal of Systems and Control Engineering.Vol. 221*, pp. 975-989, 2007.
[34] M. F. Abd Samad, H. Jamaluddin, R. Ahmad, M.S. Yaacob and A. K. M. Azad, Deterministic Mutation-Based Algorithm for Model Structure Selection in Discrete-Time System Identification. *International Journal of Intelligent Control and Systems. Vol. 16*, n. 3, pp. 182-190, 2011.
[35] X. Yao, Evolutionary Computation, In R. Sarker, , M. Mohammadian, and X. Yao, (Eds.). *Evolutionary Optimization* (Boston: Kluwer Academic Publishers., 2002, 27-56).
[36] R. Sarker, M. Mohammadian and X. Yao, (Eds.) *Evolutionary Optimization* (Boston: Kluwer Academic Publishers, 2002).
[37] T. Bäck, Introduction to Evolutionary Algorithms, In T. Bäck, D. B. Fogel and Z. Michalewicz, (Eds.) *Evolutionary Computation 1: Basic Algorithms and Operators* (Bristol: Institute of Physics Publishing, 2000, 59-63).
[38] D. B. Fogel, *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence* (Hoboken, New Jersey: John Wiley & Sons, Inc., 2006).
[39] K. A. De Jong, *Evolutionary Computation: A Unified Approach* (Cambridge, Massachusetts: The MIT Press, 2006).
[40] D. B. Fogel, Introduction to Evolutionary Computation, In T. Bäck, D. B. Fogel and Z. Michalewicz (Eds.) *Evolutionary Computation 1: Basic Algorithms and Operators* (Bristol: Institute of Physics Publishing. 2000, 1-3)
[41] J. H. Holland, Genetic Algorithms and the Optimal Allocation of Trials. *SIAM Journal of Computing. Vol. 2*, n. 2, pp. 88-105, 1973.
[42] J. H. Holland. *Adaptation in Natural and Artificial Systems* (MIT Press ed., Massachusetts Institute of Technology, 1992).
[43] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Reading, Massachusetts: Addison-Wesley Publishing Company, Inc., 1989).
[44] J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection* (Cambridge, Massachusetts: The MIT Press, 1992).
[45] J. R. Koza, Simultaneous Discovery of Detectors and a Way of Using the Detectors via Genetic Programming, *IEEE International Conference on Neural Networks.Vol. 3*, San Francisco, California, Mar 28-Apr 1, 1993, pp. 1794-1801.
[46] T. Bäck, U. Hammel and H.-P. Schwefel, Evolutionary Computation: Comments on the History and Current State, *IEEE Transactions on Evolutionary Computation. Vol. 1*, n. 1, pp. 3-17, 1997.
[47] T. Bäck, D. B. Fogel and Z. Michalewicz, (Eds.) *Evolutionary Computation 1: Basic Algorithms and Operators* (Bristol: Institute of Physics Publishing, 2000).
[48] D. Beasley, Possible Applications of Evolutionary Computation. In T. Bäck, D. B. Fogel and Z. Michalewicz (Eds.) *Evolutionary Computation 1: Basic Algorithms and Operators* (Bristol: Institute of Physics Publishing, 2000, 4-19)
[49] M. Gen and R. Cheng, *Genetic Algorithms and Engineering Design* (New York: John Wiley & Sons, Inc., 1997).
[50] A. P. Alves da Silva and P. J. Abrão, Applications of Evolutionary Computation in Electric Power Systems, *Proceedings of the 2002 Congress on Evolutionary Computation (CEC'02), Vol. 2*, Honolulu, Hawaii, USA, May 12-17, 2002, pp. 1057-1062.
[51] D. S. Weile and E. Michielssen, Genetic Algorithm Optimization Applied to Electromagnetics: A Review. *IEEE Transactions on Antennas and Propagation. Vol. 45*, n. 3, pp. 343-353, 1997.
[52] C. Dimopoulos and A. M. S. Zalzala, Recent Developments in Evolutionary Computation for Manufacturing Optimization: Problems, Solutions and Comparisons. *IEEE Transactions on Evolutionary Computation. Vol. 4*, n. 2, pp. 93-113, 2000.
[53] H. Aytug, M. Khouja, and F. E. Vergara, Use of Genetic Algorithms to Solve Production and Operations Management Problems: A Review. *International Journal of Production Research. Vol. 41*, n. 17, pp. 3955-4009. 2003.
[54] P. J. Fleming and R. C. Purshouse, Evolutionary Algorithms in Control Systems Engineering: A Survey. *Control Engineering Practice. Vol. 10*, n. 11, pp. 1223-1241. 2002.
[55] T. Bäck, Adaptive Business Intelligence Based on Evolution Strategies: Some Application Examples of Self-Adaptive Software. *Information Sciences. Vol. 148*, n. 1-4, pp. 113-121. 2002.
[56] M. S. Hossain and A. El-shafie, Intelligent Systems in Optimizing

Reservoir Operation Policy: A Review. *Water Resources Management. Vol. 27*, n. 9, pp. 3387-3407, 2013.

[57] N. Chaiyaratana and A. M. S. Zalzala, Recent Developments in Evolutionary and Genetic Algorithms: Theory and Applications. *Second IEE/IEEE International Conference on Genetic Algorithms in Engineering Systems: Innovations and Applications (GALESIA), Conference Publication No. 446*, Glasgow, UK, Sept 2-4, 1997, pp. 270-277.

[58] Al Gizi, A.J.H., Mustafa, M.W., Review of intelligent control system, (2013) *International Review of Automatic Control (IREACO)*, 6 (5), pp. 565-588.

[59] T. Kumon, M. Iwasaki, T.Suzuki, T. Hashiyama, N. Matsui and S. Okuma, Nonlinear System Identification Using Genetic Algorithms, *26th Annual Conference of the IEEE Industrial Electronics Society (IECON 2000), Vol. 4*, Nagoya, Japan, Oct 22-28, 2000, pp. 2485-2491.

[60] R. Ahmad, H. Jamaluddin and M. A. Hussain, Multivariable System Identification for Dynamic Discrete-Time Nonlinear System Using Genetic Algorithm, *IEEE International Conference on Systems, Man and Cybernetics, Vol. 5*. Hammamet, Tunisia, Oct 6-9, 2002, 6 pages.

[61] Z. Zibo and F. Naghdy, Application of Genetic Algorithms to System Identification, *Proceedings of the 1995 International Conference on Evolutionary Computing, Vol. 2,* Perth, Western Australia, Nov 29-Dec 1, 1995, pp. 777-782.

[62] V. Duong and A. R. Stubberud, System Identification by Genetic Algorithm, *IEEE Aerospace Conference Proceedings, Vol. 5*, Big Sky, Montana, Mar 9-16, 2002, pp. 2331-2337.

[63] K. C. Tan, Y. Li, D. J. Murray-Smith and K. C. Sharman, System Identification and Linearisation Using Genetic Algorithms with Simulated Annealing, *First IEE/IEEE International Conference on Genetic Algorithms in Engineering Systems: Innovations and Applications (GALESIA), Conference Publication No. 414*, Sheffield, UK, Sept 12-14, 1995, pp. 164-169.

[64] A. B. Tcholakian, A. Martins, R. C. S. Pacheco and R. M. Barcia, Fuzzy System Identification Through Hybrid Genetic Algorithm, *1997 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS'97)*, Syracuse, New York, Sept 21-24, 1997, pp. 428-433.

[65] K. Kristinsson and G. A. Dumont, System Identification and Control Using Genetic Algorithms. *IEEE Transactions on Systems, Man and Cybernetics. Vol. 22*, n. 5, pp. 1033-1046, 1992.

[66] T. Hatanaka, K. Uosaki and M. Koga, Evolutionary Computation Approach to Block Oriented Nonlinear Model Identification, *5th Asian Control Conference*. *Vol. 1*, Melbourne, Australia, July 20-23, 2004, pp. 90-96.

[67] X. Tan and H. Yang, The Optimization of Nonlinear Systems Identification Based on Genetic Algorithms, *International Conference on Computational Intelligence and Security, Vol. 1*, Guangzhou, China, Nov 3-6, 2006, pp. 266-269.

[68] J.-G. Juang, Application of Genetic Algorithm and Recurrent Network to Nonlinear System Identification, *Proceedings of 2003 IEEE Conference on Control Applications, Vol. 1,* Istanbul, Turkey, June 23-25, 2003, pp. 129-134.

[69] H. M. Abbas and M. M. Bayoumi, Volterra-System Identification Using Adaptive Real-Coded Genetic Algorithm. *IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans. Vol. 36*, n. 4, pp. 671-684, 2006.

[70] M. Iwasaki, M. Miwa and N. Matsui, GA-Based Evolutionary Identification Algorithm for Unknown Structured Mechatronic Systems. *IEEE Transactions on Industrial Electronics*. *Vol. 52*, n. 1, pp. 300-305, 2005.

[71] M. Rocha, P. Sousa, P. Cortez and M. Rio, Quality of Service Constrained Routing Optimization using Evolutionary Computation. *Applied Soft Computing, Vol. 11*, n. 1, pp. 356–364, 2011.

[72] O. Valenzuela, B. Delgado-Marquez and M. Pasadas, Evolutionary Computation for Optimal Knots Allocation in Smoothing Splines, *Applied Mathematical Modelling. Vol. 37*, n. 8, pp. 5851–5863, 2013.

[73] J. F. Izquierdo and J. Rubio, Antenna Modeling by Elementary Sources Based on Spherical Waves Translation and Evolutionary Computation. *IEEE Antennas and Wireless Propagation Letters. Vol. 10*, pp. 923 – 926, 2011.

[74] D. Marco, D. Cairns and C. Shankland, Optimisation of Process Algebra Models using Evolutionary Computation, *IEEE Congress on Evolutionary Computation (CEC)*, New Orleans, Louisiana, June 5-8, 2011, pp. 1296 – 1301.

[75] V. K. Dertimanis, D. V. Koulocheris and C. N. Spentzas, Design of a Hybrid Algorithm for ARMA Parameter Estimation. *5th GRACM International Congress on Computational Mechanics*, *Vol. 1*, Limassol, Cyprus, June 29-July 1, 2005, pp. 395-402.

[76] C. Unsihuay-Vila, A. C. Zambroni de Souza, J. W. Marangon-Lima and P. P. Balestrassi, Electricity Demand and Spot Price Forecasting using Evolutionary Computation Combined with Chaotic Nonlinear Dynamic Model. *International Journal of Electrical Power & Energy Systems*. *Vol. 32*, n. 2, pp. 108–116, 2010.

[77] C. Sales, R. M. Rodrigues, F. Lindqvist, J. Costa, A. Klautau, K. Ericson, J. Rius i Riu and P. O. Borjesson, Line Topology Identification using Multiobjective Evolutionary Computation. *IEEE Transactions on Instrumentation and Measurement. Vol. 59*, n. 3, pp. 715-729, 2010.

[78] D. Koulocheris, V. Dertimanis and H. Vrazopoulos, Evolutionary Parametric Identification of Dynamic Systems. *Forschung im Ingenieurwesen. Vol. 68*, n. 4, pp. 173-181, 2004.

[79] D. S. Pereira and J. O. P. Pinto, Genetic Algorithm Based System Identification and PID Tuning for Optimum Adaptive Control, *Proceedings of the 2005 IEEE/ASME International Conference on Advanced Intelligent Mechatronics.* Monterey, California, July 24-28, 2005, pp. 801-806.

[80] M. J. Perry, G. C. Koh and Y. S. Choo, Modified Genetic Algorithm Approach to Structural Identification, *The 3rd International Conference on Structural Stability and Dynamics*, Kissimmee, Florida, June 19-22, 2005, 6 pages.

[81] K. Asan Mohideen, G. Saravanakumar, K. Valarmathi, D. Devaraj and T.K. Radhakrishnan, Real-Coded Genetic Algorithm for System Identification and Tuning of a Modified Model Reference Adaptive Controller for a Hybrid Tank System. *Applied Mathematical Modelling. Vol. 37*, n. 6, pp. 3829–3847, 2013.

[82] A. Kusiak, H. Zheng, and Z. Song, Power optimization of wind turbines with data mining and evolutionary computation. *Renewable Energy. Vol. 35*, n. 3, pp. 695–702, 2010.

[83] S. Salcedo-Sanz, D. Gallo-Marazuela, A. Pastor-Sánchez, L. Carro-Calvo, A. Portilla-Figueras and L. Prieto, Evolutionary computation approaches for real offshore wind farm layout: A case study in northern Europe. *Expert Systems with Applications. Vol. 40*, n. 16, pp. 6292–6297, 2013.

[84] A. H. Wright, Genetic Algorithms for Real Parameter Optimization. In G. J. E. Rawlins, (Ed.) *Foundations of Genetic Algorithms* (San Mateo, California: Morgan Kaufmann Publishers, Inc., 1991, 205-220)

[85] T. Bäck and D. B. Fogel, Glossary, In T. Bäck, D. B. Fogel and Z. Michalewicz, (Eds.). *Evolutionary Computation 1: Basic Algorithms and Operators* (Bristol: Institute of Physics Publishing, 2000, xxi-xxxvii)

[86] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs* (3rd, revised and extended ed., Berlin: Springer-Verlag, 1996).

[87] A. Czarn, C. MacNish, K. Vijayan and B. Turlach, Statistical Exploratory Analysis of Genetic Algorithms: The Importance of Interaction, *Proceedings of the 2004 IEEE Congress on Evolutionary Computation (CEC'04), Vol. 2,* Portland, Oregon, June 19-23, 2004, pp. 2288-2295.

[88] A. M. S. Zalzala and P. J. Fleming (Eds.) *Genetic Algorithms in Engineering Systems* (London, United Kingdom: The Institution of Electrical Engineers, 1997).

[89] T. A. Johansen, Identification of Non-Linear Systems using Empirical Data and Prior Knowledge – An Optimization Approach. *Automatica. Vol. 32*, n. 3, pp. 337-356, 1996.

[90] Y.-X. Li and M.Gen, Nonlinear Mixed Integer Programming Problem Using Genetic Algorithm and Penalty Function, *IEEE International Conference on Systems, Man and Cybernetics, Vol. 4*, Beijing, China, Oct 14-17, 1996, pp. 2677-2682.

[91] T. Yamada and C. R. Reeves, Permutation Flowshop Scheduling by Genetic Local Search. *Second IEE/IEEE International Conference on Genetic Algorithms in Engineering Systems:*

Innovations and Applications (GALESIA), Conference Publication No. 446, Glasgow, UK, Sept 2-4, 1997, pp. 270-277.
[92] A. Kanarachos, D. Koulocheris and H. Vrazopoulos, Evolutionary Algorithms with Deterministic Mutation Operators Used for the Optimization of the Trajectory of a Four-Bar Mechanism. *Mathematics and Computers in Simulation. Vol. 63*, n. 6, pp. 483-492, 2003.
[93] Y. Liang and K. S. Leung, Two-Way Mutation Evolution Strategies, *Proceedings of the 2002 IEEE Congress on Evolutionary Computation (CEC'02), Vol. 1,* Honolulu, Hawaii May 12-17, 2002, pp. 789-794.
[94] J. J. Grefenstette, Optimization of Control Parameters for Genetic Algorithms. *IEEE Transactions on Systems, Man and Cybernetics. Vol. 16*, n. 1, pp. 122-128, 1986.
[95] H. Mühlenbein, L. Zinchenko, V. Kureichik and T. Mahnig, Effective Mutation Rate for Probabilistic Evolutionary Design of Analogue Electrical Circuits. *Applied Soft Computing. Vol. 7*, n. 3, pp. 1012-1018, 2007.
[96] K. A. De Jong, *An Analysis of the Behavior of a Class of Genetic Adaptive Systems*. Ph.D. dissertation, The University of Michigan, 1975.
[97] L. B. Booker, D. B. Fogel, D. Whitley, P. J. Angeline and A. E. Eiben, Recombination. In T. Bäck, D. B. Fogel and Z. Michalewicz, (Eds.) *Evolutionary Computation 1: Basic Algorithms and Operators* (Bristol: Institute of Physics Publishing, 2000, 256-307)
[98] T. Murata and H. Ishibuchi, Positive and Negative Combination Effects of Crossover and Mutation Operators in Sequencing Problems, *Proceedings of the 1996 IEEE International Conference on Evolutionary Computation*, Nagoya, Japan, May 20-22, 1996, pp. 170-175.
[99] A. E. Eiben, R. Hinterding and Z. Michalewicz, Parameter Control, In T, Bäck, D. B. Fogel, and Z. Michalewicz (Eds.) *Evolutionary Computation 2: Advanced Algorithms and Operators* (Bristol: Institute of Physics Publishing, 2000, 170-187).

Authors' information

Md Fahmi Abd Samad received his Bachelor in Mechanical Engineering, Masters in Engineering Management and Ph.D. in Mechanical Engineering from Universiti Teknologi Malaysia, Johore, Malaysia in 2000, 2002 and 2009, respectively. His interests are in system identification, evolutionary computation and control engineering. Currently, he works as a Senior Lecturer in Faculty of Mechanical Engineering, Universiti Teknikal Malaysia Melaka. Dr Md Fahmi Abd Samad is a Graduate Member of Institution of Engineers, Malaysia.

