

**THE PRE-PROCESSING TECHNIQUE AND PARAMETER ADJUSTMENT
INFLUENCES ON NARX-BASED BPSO STRUCTURE FOR TIME-VARYING
WATER TEMPERATURE MODELLING**

N. Hambali, M. N. Taib, I. M. Yassin and M. H. F. Rahiman*

Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor,
Malaysia

Published online: 17 October 2017

ABSTRACT

This study investigates the effects of pre-processing technique and parameter adjustment on the fitness by utilizing a Binary Particle Swarm Optimization (BPSO) algorithm. The model structure selection of polynomial NARX has been focused on BPSO algorithm for time-varying water temperature modelling of Steam Distillation Pilot Plant (SDPP). Magnitude scaling, block division and interleaving were employed for pre-processing technique while various swarm sizes, maximum iterations, and random seeds parameters have been combined for BPSO parameter adjustment. The combination of interleaving and magnitude scaling techniques demonstrate the best minimum correlation violations (CRV) and fitness. Subsequently, higher swarm size slightly improved while the maximum iteration and initial random seed alterations had an insubstantial consequence on fitness values for the parameter adjustment.

Keywords: identification; NARX; particle swarm optimization; distillation column; temperature.

Author Correspondence, e-mail: hezrif@ieee.org

doi: <http://dx.doi.org/10.4314/jfas.v9i5s.23>



1. INTRODUCTION

The nonlinear dynamic behaviour has been displayed by a broad application of the distillation column in chemical operations [1-2]. Auto-Regressive with eXogenous input (ARX) [3-4] has been used for distillation column as well as Auto-Regressive Moving Average with eXogenous input (ARMAX) [5] and Nonlinear Auto-Regressive with eXogenous input (NARX) [6-7].

Despite this, some restriction regarding the robustness of particular input range [8], less accuracy and poor fit has been reported for a nonlinear system using linear modelling compared to the nonlinear model which offers more beneficial operation, potentiality and tolerable for the nonlinear system representation [4, 9].

So far, however, there has been little discussion about NARX for the steam distillation column in [5, 10]. In addition, no research has been found that applied Binary Particle Swarm Optimization (BPSO) algorithm for the mentioned distillation column above using polynomial NARX model.

The BPSO is an altered of the original PSO algorithm for binary optimization problem solving [11]. The particle values [7] and discretized version [12] for binary problems solution have been employed using a PSO algorithm. The proposed BPSO method was not only demonstrated higher precision [13-15], but also tightened power loss and competent energy management [16-17]. BPSO was skilful in the model fit enhancement and correlation violations (CRV) number decline [18]. Moreover, previous studies have reported on the BPSO parameter adjustment effects. Selected swarm size has been influenced the DC Motor structure selection optimization with good fitness values [18-19].

This paper has been divided into six parts. The second part deals with the theoretical background. In the third part, the experimental setup is presented. The methodology that has been used in this paper is also reported in part four. All the results and discussion has been shown in part five. Lastly, the conclusion and future work are discussed in the final part.

2. SYSTEM IDENTIFICATION

The NARX is an extension of ARX and does not reuse residual terms back into the model like NARMAX. The NARX model is presented as

$$y(t) = f^d \left[\begin{array}{c} y(t-1), y(t-2), \dots, y(t-n_y), \\ u(t-n_k), u(t-n_k-1), \dots, u(t-n_k-n_u) \end{array} \right] + \varepsilon(t) \quad (1)$$

f^d is the estimated model, $y(t)$, $u(t)$ are the output and input respectively and n_y and n_u are their corresponding maximum lags. Parameter n_k is the input signal time delay with typical value 1, otherwise $n_k = 0$ when the input $u(t)$ is needed for identification. The pick of model structures demands choosing which lagged y , u and ε terms and forecasting the parameters that can describe the upcoming values of y .

2.1. Polynomial NARX

NARX polynomial model representation [20] is given by

$$y(t) = \sum_{m=1}^{n_p} P_m \theta_m + \varepsilon(t) \quad (2)$$

P_m is the m -th regression term with $P_1 = 1$, θ_m is the m -th regression parameter and n_p is the number of terms in the polynomial expansion. The formulation and solution of Least Squares (LS) problem have been involved in identification.

$$P\theta + \varepsilon = y \quad (3)$$

y is the $n \times 1$ vector of real reflections, θ is a $m \times 1$ coefficient vector and P is a $n \times m$ regressor matrix. The PSO technique is established on evolutionary computation and swarm philosophy. The convergence quality improvement and the algorithm adaptation in problems solving contributed to numerous established variants such as Vanilla and Binary PSO algorithms.

2.2. Model Structure Selection: Binary Particle Swarm Optimization (BPSO)

The velocity and position update equations are among the Vanilla PSO algorithm;

$$V_{id} = V_{id} + C_1(P_{best} - X_{id}) \times rand_1 + C_2(G_{best} - X_{id}) \times rand_2 \quad (4)$$

The value of V_{id} has been employed for a particle positions alteration.

$$X_{id} = X_{id} + V_{id} \quad (5)$$

P_{best} and G_{best} are the best particle's fitness and solution achieved by the swarm respectively. C_1 and C_2 are the cognition and social learning rate respectively. While V_{id} and X_{id} are the particle velocity and position respectively. $rand_1$ and $rand_2$ are uniformly-distributed random

numbers between 0 and 1.

The probabilities of change have been demonstrated in the BPSO particle positions rather than the actual solution. The bit change process is;

$$\text{bin string} = \begin{cases} 1, & X_{id} \geq 0.5 \\ 0, & X_{id} < 0.5 \end{cases} \quad (6)$$

The particle positions (probabilities) are between 0 and 1. A probability value greater than 0.5 specifies that the bit will vary from its current condition to another (either 0 to 1 or 1 to 0). Else, if the particle value is lesser than 0.5, the bit will sustain.

In the swarm of BPSO for polynomial structure selection, a $1 \times m$ solutions vector, X_{id} has been transmitted for each one particle. QR factorization has been employed for the forecasting of the parameter value, θ_R for the reduced P matrix (θ_R).

$$P_R \theta_R + \varepsilon = y \quad (7)$$

$$P_R = Q_R R_R \quad (8)$$

$$g_R = Q_R^T y \quad (9)$$

$$R_R \theta_R = g_R \quad (10)$$

Next, the value of θ_R can be estimated by reorganizing and solving (10)[20].

$$\theta_R = R_R^T g_R \quad (11)$$

2.2 Model Estimation

Based on the model parameters, θ , the residuals Normalised Sum Squared Error (NSSE), $V_{NSSE}(\theta, Z^N)$ is;

$$V_{NSSE}(\theta, Z^N) = \frac{1}{2N} \sum_{t=1}^N \varepsilon^2(t, \theta) \quad (12)$$

The selection of model order can be done by using several established models selection criteria such as Akaike Information Criterion (AIC), Model Descriptor Length (MDL) and Final Prediction Error (FPE) as shown in (13), (14) and (15) respectively.

$$V_{AIC} = \left(1 + 2 \frac{d}{N}\right) V_{NSSE}(\theta, Z^N) \quad (13)$$

$$V_{MDL} = \left(1 + \log(N) \frac{d}{N}\right) V_{NSSE}(\theta, Z^N) \quad (14)$$

$$V_{FPE} = \left(\frac{1 + \frac{d}{N}}{1 - \frac{d}{N}}\right) V_{NSSE}(\theta, Z^N) \quad (15)$$

where the number of estimated parameters and the data points amount are presented as d and N , respectively. Models with the lowest V_{AIC} , V_{MDL} and V_{FPE} scores conform the principle of parsimony as the smallest quantity of parameters were necessary to supply the most beneficial fit for the data.

The collected data using SDPP will be separated into two divisions; training set for model estimation and testing set for model validation. For model estimation, the sample data are utilized for estimation to fit the NARX model. Three pre-processing (PP) methods namely magnitude scaling, block division, and interlacing are presented.

Per a standard range, the magnitude scaling method will scale the data. The scaling technique is required when there is an unsuitable input-output data sets range. The scaling equation is given by;

$$y = \frac{(y_{max}-y_{min})(x_{max}-x_{min})}{x_{max}-x_{min}} + y_{min} \quad (16)$$

$(y_{max} - y_{min})$ and $(x_{max} - x_{min})$ are the data range after and prior to scaling respectively. y and x are the data before and rescaled, respectively.

Typically, the equal 50 % division is for training and testing. Two methods of division exist, namely block division and interlacing. Only one of these methods is used at one time. In block division, the first n cases dataset is assign to the training set ($D_{training}$) and the leftover dataset for the testing set ($D_{testing}$).

The interlacing method divides the dataset based on the position of the data into training and testing sets. The training set contains odd positions data, while even positions data are allotted for the testing set.

For odd N case, $D_{1,3,5,\dots,N} = D_{training}$ and $D_{2,4,6,\dots,N-1} = D_{testing}$.

Otherwise, for even N case, $D_{2,4,6,\dots,N-1} = D_{training}$ and $D_{1,3,5,\dots,N} = D_{testing}$.

3. EXPERIMENTAL DESIGN

A coil-type heater immersed in water for SDPP system is employed to generate steam [10]. The usage of two resistive temperature detectors (RTD) PT-100 is applied to monitor the water temperature and the steam in the column. The resistance output from both sensors is

converted within 1 V to 5 V for certain various temperature ranges. A signal modulates a 1.5 kW, 240 V and 50 Hz power carried to submersion heating component with specific sampling time. The MATLAB receives all the samples of the process. Fig. 1 illustrates the SDPP system.

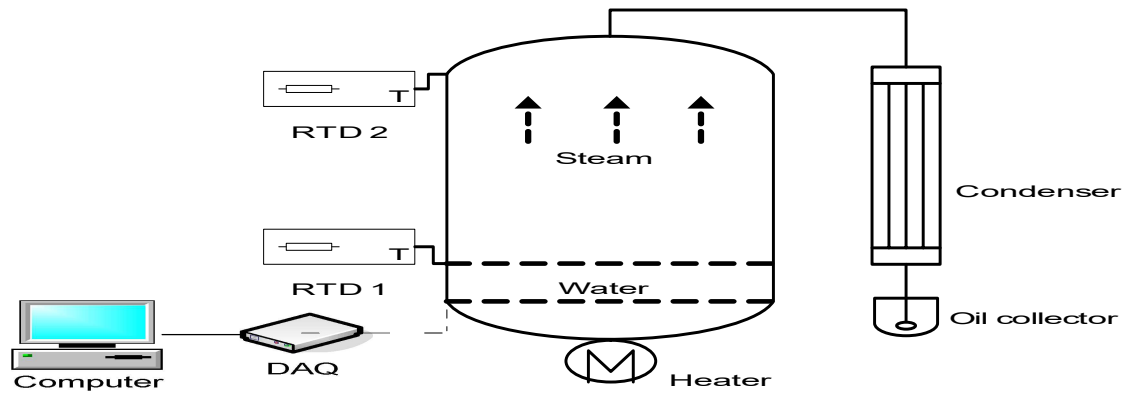


Fig.1. Steam distillation pilot plant (SDPP) system

4. METHODOLOGY

The nonlinear water temperature dataset that consists of 18, 000 data points from the SDPP system with the implementation of the PRBS input is demonstrated in Fig. 2.

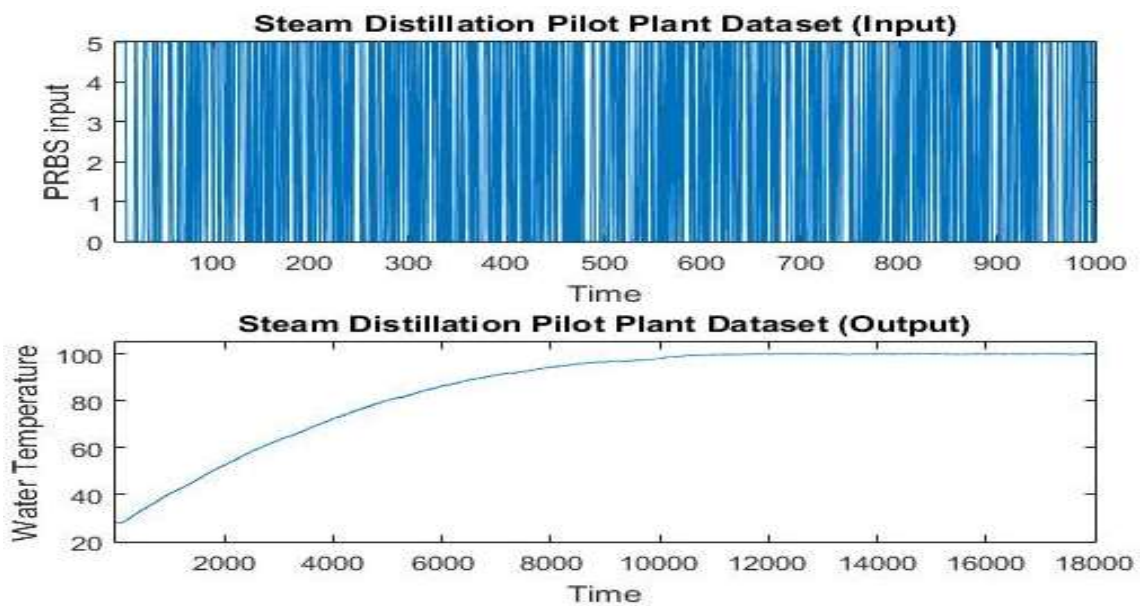


Fig.2. SDPP dataset

A total of four PP combinations were utilized in this study. The methods are shown in Table 1.

Table 1.Pre-processing technique and respective codes

PP Code	Pre-Processing Method
00	No magnitude scaling, block division
01	No magnitude scaling, interleaving
10	Magnitude scaling, block division
11	Magnitude scaling, interleaving

For optimization purposes, various swarm sizes, maximum iterations and random seeds parameters have been combined. Higher swarm sizes create larger potential in global minima searching by the number of agents based on optimization time and computational cost. Termination of the PSO search will be driven by the achievement of the objective or the finding of maximum iterations for each experiment. In addition, the initial random seed for each V_{id} and X_{id} particles have influence on the final optimisation PSO result. The results consistency will be decided by different initial random seeds for repeated experiments. Therefore, for the fitness function optimal convergence searching, the selected parameter values as listed in Table 2 are tested as been implemented by [18-19].

Table 2. BPSO parameter settings for structure selection

Parameter	Value
Fitness Criterion	AIC, FPE, MDL
Swarm size	10, 20, 30, 40, 50
Maximum Iterations	500, 1000, 1500
Initial Random Seed	0, 10 000, 20 000
X_{min}	0
X_{max}	1
V_{min}	-1
V_{max}	+1
C_1	2.0
C_2	2.0

5. RESULTS AND DISCUSSION

An analysis of frequency term selection was performed for the NARX water temperature model for all PP. The objectives of the test were to determine whether the BPSO algorithm considered all terms to be included in the model, and whether any selection patterns were present that indicate a preference towards certain terms that represent the system well. The term selection frequencies for the NARX water temperature models are shown in the next figures.

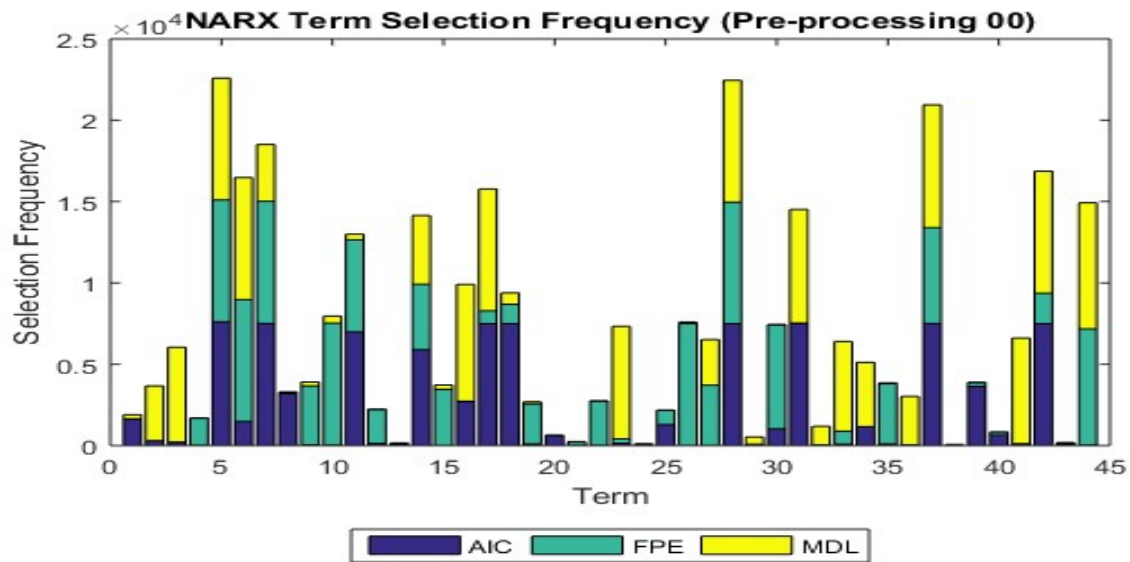


Fig.3. NARX term selection frequency (PP 00)

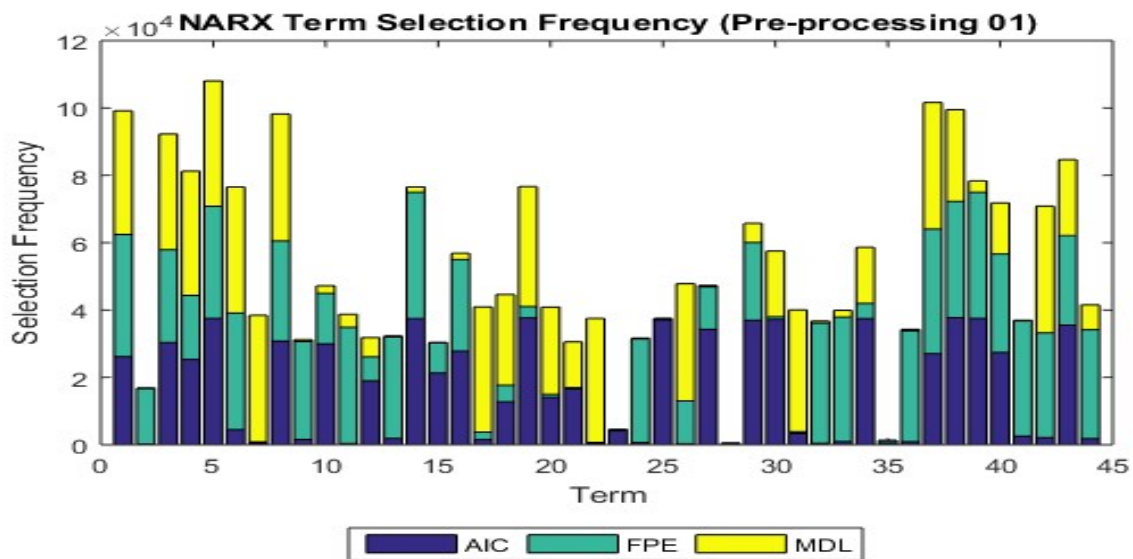


Fig.4. NARX term selection frequency (PP 01)

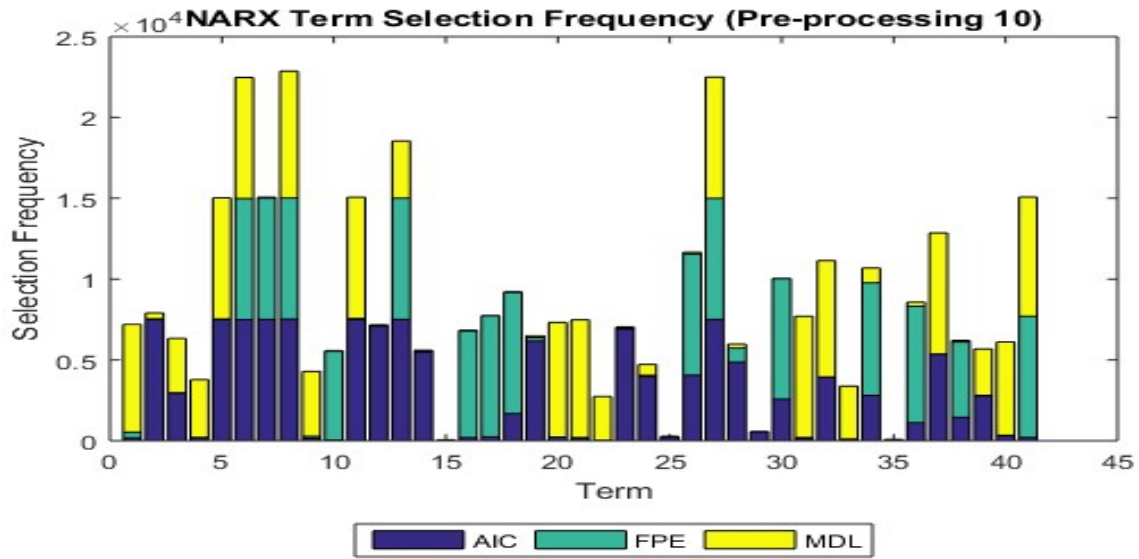


Fig.5. NARX term selection frequency (PP 10)

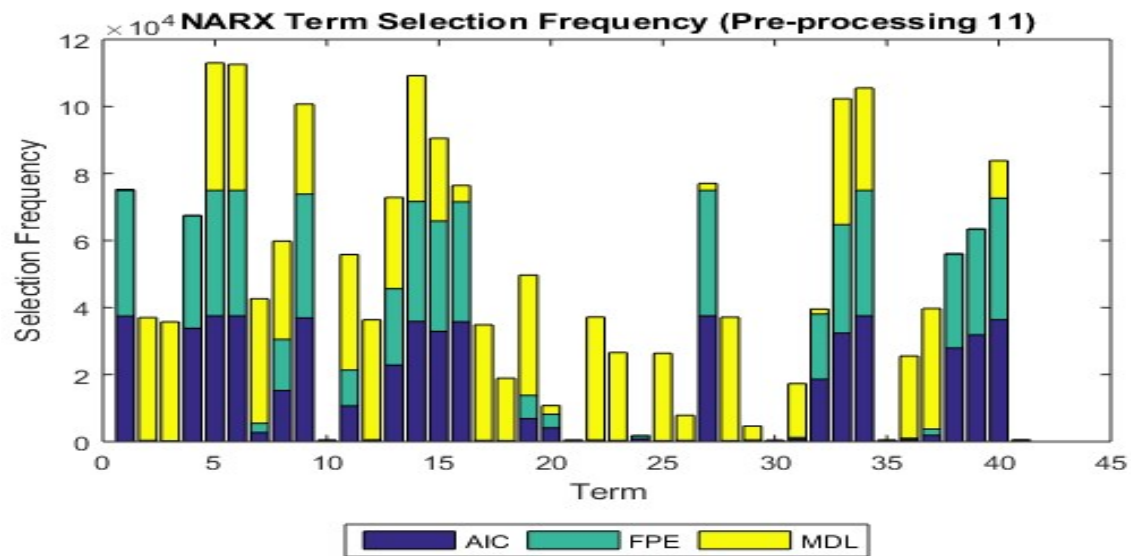


Fig.6. NARX term selection frequency (PP 11)

The term selection frequency for all PP indicates that BPSO considered all terms during optimization, as most all terms are selected at least once by the BPSO algorithm for all models selection criteria, AIC, FPE and MDL. Moreover, there appear to be random searches on both low-performing and high-performing terms.

The water temperature dataset has input and output lag space of four ($n_u = 4, n_y = 4$). It can be seen that the magnitude scaling technique for PP 10 and PP 11 has only 41 terms as shown in Fig. 5 and Fig. 6 compared to PP 00 and PP 01 (without magnitude scaling) which have 44

terms (Fig. 3 and Fig. 4).

Table 3 indicates the fitness and total CRV of training and testing values based on the minimum fitness searching. The magnitude scaling and block division techniques from PP 11 resulted in the lowest fitness. On the other side, the magnitude scaling resulted in higher CRV for PP 10 and 11 which fails the correlation tests.

Table 3. Effect of pre-processing method on fitness and total CRV based on minimum fitness search

PP Code	Criterion	Fitness Value			Total CRV
		Min	Average	Max	
00	AIC	5.4905×10^{-3}	5.9162×10^{-3}	6.3549×10^{-3}	65
	FPE	5.5696×10^{-3}	5.9629×10^{-3}	6.6241×10^{-3}	140
	MDL	5.3399×10^{-3}	5.6995×10^{-3}	6.1529×10^{-3}	65
01	AIC	5.7183×10^{-3}	6.0288×10^{-3}	6.4284×10^{-3}	52
	FPE	5.8063×10^{-3}	6.1875×10^{-3}	6.7334×10^{-3}	52
	MDL	5.5615×10^{-3}	5.8595×10^{-3}	6.2266×10^{-3}	52
10	AIC	4.3213×10^{-6}	4.6457×10^{-6}	5.2024×10^{-6}	95
	FPE	4.3975×10^{-6}	4.9249×10^{-6}	6.0915×10^{-6}	95
	MDL	4.1847×10^{-6}	4.5497×10^{-6}	4.8951×10^{-6}	95
11	AIC	4.5570×10^{-6}	4.8525×10^{-6}	5.6715×10^{-6}	181
	FPE	4.6021×10^{-6}	4.9651×10^{-6}	6.1327×10^{-6}	174
	MDL	4.3892×10^{-6}	4.6071×10^{-6}	5.2881×10^{-6}	129

Complementary to this, based on the minimum CRV searching as illustrated in Table 4, interleaving without magnitude scaling techniques recorded the best combination of both CRV and fitness from PP 01. In this search, the NARX-based BPSO model managed to cut down the CRV numbers and maintained the low fitness values in PP 01 while other PP methods demonstrated high CRV numbers. Therefore, the best results were obtained using PP method 01. This is based on the lowest total CRV from both training and testing sets and low fitness value. In addition, lower CRV values contribute to the uncorrelated residuals which represent a good model fit.

Table4. Effect of pre-processing method on fitness and total CRV based on minimum CRV search

PP Code	Criterion	Min Total CRV	Fitness Value
00	AIC	61	5.9268×10^{-3}
	FPE	37	5.7732×10^{-3}
	MDL	62	5.5275×10^{-3}
01	AIC	9	6.1741×10^{-3}
	FPE	11	6.7334×10^{-3}
	MDL	10	5.7969×10^{-3}
10	AIC	47	5.0208×10^{-6}
	FPE	95	4.3975×10^{-6}
	MDL	54	4.8866×10^{-6}
11	AIC	52	4.7407×10^{-6}
	FPE	117	4.9478×10^{-6}
	MDL	51	4.5265×10^{-6}

Three parameters; swarm size, maximum iterations and initial random seed as listed in Table 2 influenced the BPSO convergence. The parameters consequence on convergence is analyzed for pre-processing method 01. Swarm sizes between 10 and 50 presented low fitness variations as evident from Fig. 7. Furthermore, it can be observed that swarm size 50 showed marginally better fitness values. Whereas, as can be found in Fig. 8 and Fig. 9, the maximum iterations parameter and the random seed respectively, indicated insignificant outcomes on the fitness. Moreover, a small number of iterations which is lesser than 15 in Fig. 10 was reported for convergence of iteration. This is predicted since dissimilar initialization seeds lead contrarily to the exploration progression of the solution space [18].

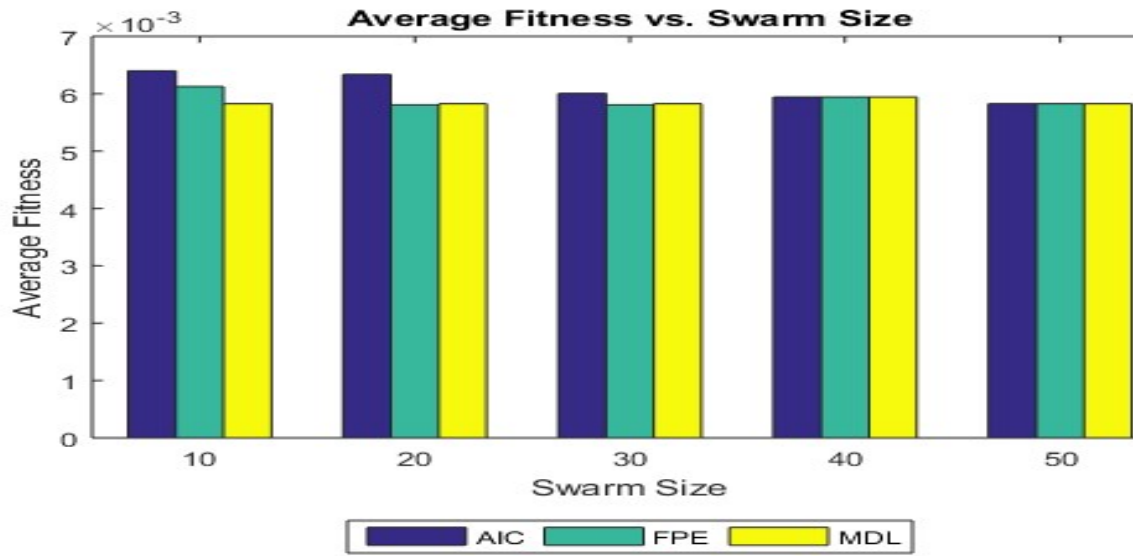


Fig.7. Average fitness vs. swarm size

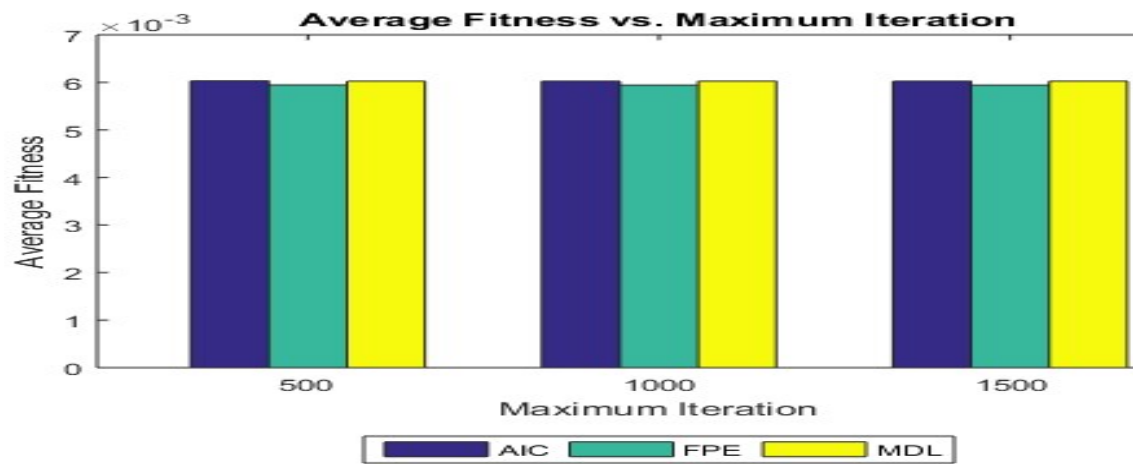


Fig.8. Average fitness vs. maximum iteration

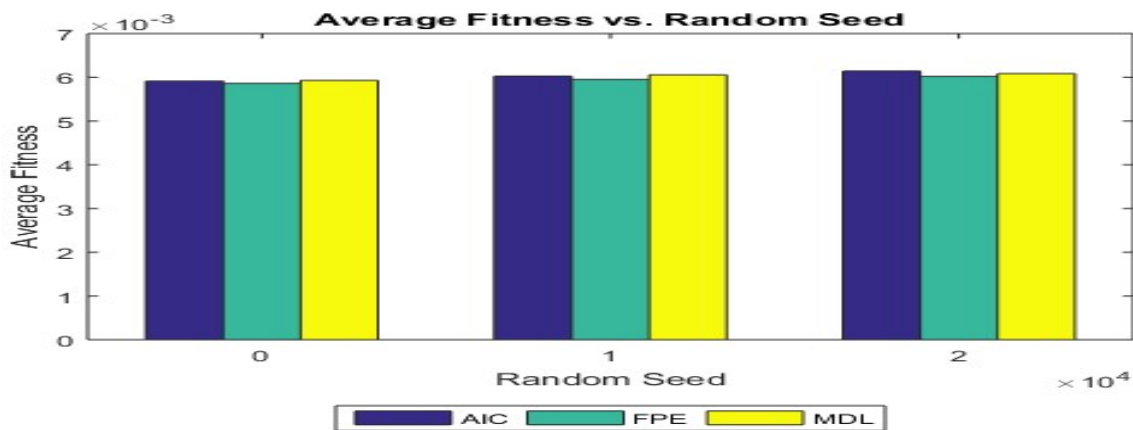


Fig.9. Average fitness vs. random seed

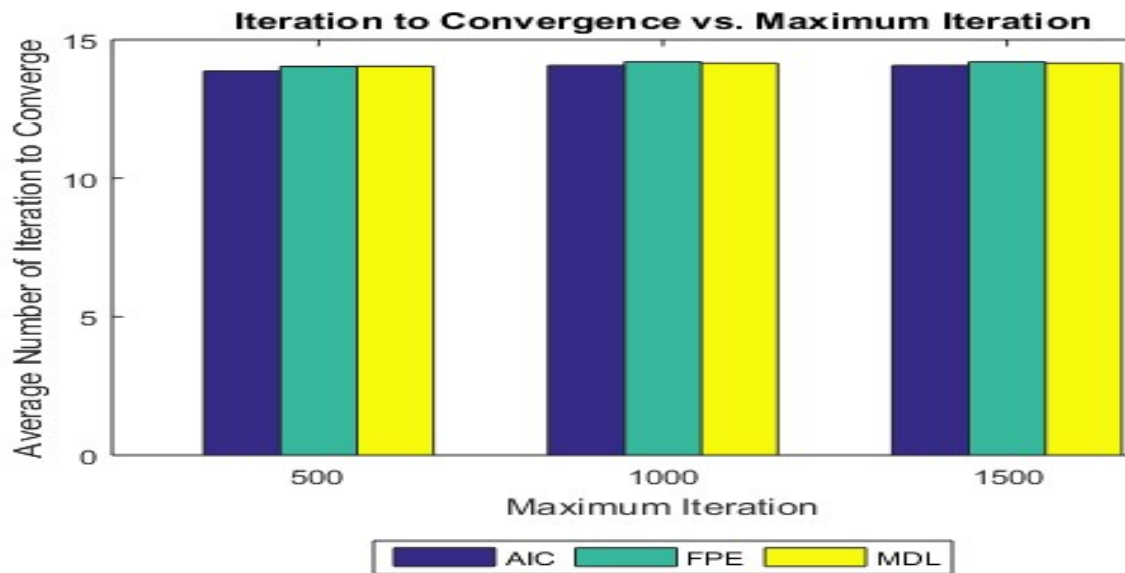


Fig.10. Iteration to convergence vs. maximum iteration

6. CONCLUSION

The effects of pre-processing technique and parameter adjustment on the fitness were identified. The system identification of SDPP employed a model structure selection of polynomial NARX that concentrated on BPSO algorithm. The interleaving method without magnitude scaling (PP 01) demonstrates the best minimum CRV and fitness, due to the minimum CRV exploration. Subsequently, for the parameter adjustment, higher swarm size provide slightly better fitness values. Likewise, there is an insubstantial consequence on the fitness with the maximum iterations and initial random seed alterations. An important question for future studies is the exploration of the uncorrelated residuals performance of the model. This optimization finding is promising and should be explored with another type of dataset, for instance, the steam temperature dataset that can be collected using SDPP too.

7. ACKNOWLEDGEMENTS

This project was financially supported by Faculty of Electrical Engineering, UiTM Shah Alam, and the SLAB/SLAIKementerianPendidikanTinggi Malaysia scholarship to the first author.

8. REFERENCES

- [1] Dudret S, Beauchard K. Stability and asymptotic observers of binary distillation processes described by nonlinear convection/diffusion models. In IEEE American Control Conference, 2012, pp. 3352-3358
- [2] Bachnas A. Linear parameter-varying modelling of a high-purity distillation column. Master thesis, Netherlands: Delft University of Technology, 2012
- [3] Najib M S, Ahmad M U, Funk P, Taib M N. Agarwood classification: A case-based reasoning approach based on e-nose. In 8th IEEE International Colloquium on Signal Processing and its Applications, 2012, pp. 120-126
- [4] Araromi D O, Sonibare J A, Emuoyibofarhe J O. Fuzzy identification of reactive distillation for acetic acid recovery from waste water. Journal of Environmental Chemical Engineering, 2014, 2(3):1394-1403
- [5] Mohammad N N, Kasuan N, Rahiman M H F, Taib M N. Steam temperature control using fuzzy logic for steam distillation essential oil extraction process. In IEEE Control and System Graduate Research Colloquium, 2011, pp. 53-58
- [6] Nordin M N N, Rahiman M H F, Adnan R, Yusoff Z M, Yassin I M. Optimizations of NARX lag space selection for a multi-layer perceptron (MLP)-based model of a down-flowing steam distillation system using particle swarm optimization (PSO). In 8th IEEE International Colloquium on Signal Processing and its Applications, 2012, pp. 533-538
- [7] Awadz F, Yassin I M, Rahiman M H F, Taib M N, Zabidi A, Hassan H A. System identification of essential oil extraction system using non-linear autoregressive model with Exogenous Inputs (NARX). In IEEE Control and System Graduate Research Colloquium, 2010, pp. 20-25
- [8] Paduart J, Lauwers L, Swevers J, Smolders K, Schoukens J, Pintelon R. Identification of nonlinear systems using polynomial nonlinear state space models. Automatica, 2010, 46(4):647-656
- [9] Shariff H M, Marzaki M H, Tajjudin M, Rahiman M H F. System identification for steam distillation pilot plant: Comparison between linear and nonlinear models. In 3rd IEEE International Conference on System Engineering and Technology, 2013, pp. 263-268

-
- [10] Kasuan N, Taib M N, Rahiman M H F. Model reference adaptive controller to regulate steam temperature in distillation process for essential oil extraction. In 7th IEEE International Colloquium on Signal Processing and its Applications, 2011, pp. 298-303
- [11] Kennedy J, Eberhart R C. A discrete binary version of the particle swarm algorithm. In IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation, 1997, pp. 4104-4108
- [12] Liu Y, Zhang Z, Liu Z. Customized configuration for hierarchical products: Component clustering and optimization with PSO. International Journal of Advanced Manufacturing Technology, 2011, 57(9):1223-1233
- [13] Keshavarzi R, Akhlaghi M, Emami F. Binary PSO algorithm assisted to investigate the optical sensor based plasmonic nano-bi-domes. Optik-International Journal for Light and Electron Optics, 2016, 127(19):7670-7675
- [14] Haixiang G, Yijing L, Yanan L, Xiao L, Jinling L. BPSO-adaboost-KNN ensemble learning algorithm for multi-class imbalanced data classification. Engineering Applications of Artificial Intelligence, 2016, 49:176-193
- [15] Beheshti Z, Shamsuddin S M, Hasan S. Memetic binary particle swarm optimization for discrete optimization problems. Information Sciences, 2015, 299:58-84
- [16] Pookpant S, Ongsakul W. Design of optimal wind farm configuration using a binary particle swarm optimization at Huasai district, Southern Thailand. Energy Conversion and Management, 2016, 108:160-180
- [17] Phuangpornpitak N, Tia S. Optimal photovoltaic placement by self-organizing hierarchical binary particle swarm optimization in distribution systems. Energy Procedia, 2016, 89:69-77
- [18] Yassin IM, Taib M N, Adnan R. Extended analysis of BPSO structure selection of nonlinear auto-regressive model with exogenous inputs (NARX) of direct current motor. Songklanakarin Journal of Science and Technology, 2014, 36(6):683-699
- [19] Yassin I M, Taib M N, Adnan R, Salleh M K M, Hamzah M K. Effect of swarm size parameter on binary particle swarm optimization-based NARX structure selection. In IEEE Symposium on Industrial Electronics and Applications, 2012, pp. 219-223

[20] Amisigo B, Van de Giesen N. Monthly streamflow prediction in the Volta Basin of West Africa: A SISO NARMAX polynomial modelling. *Physics and Chemistry of the Earth, Parts A/B/C*, 2008, 33(1):141-150

How to cite this article:

Hambali N, Taib M N, Yassin I M, Rahiman M H F. The pre-processing technique and parameter adjustment influences on narx-based bpsso structure for time-varying water temperature modelling. *J. Fundam. Appl. Sci.*, 2017, 9(5S), 320-335.