

System Identification and Model Validation of Recursive Least Squares Algorithm for Box–Jenkins Systems

Nasar Aldian Ambark Shashoa
Electrical and Electronics Engineering Department
Azzaytuna University
Tarhuna, Libya
dr.naser.elec@gmail.com

Ibrahim N. Jleta
Department of Electrical Engineering
Libyan Academy of Graduate Studies
Tripoli, Libya

Abdurrezag S. Elmezughi
Computer Engineering Department
Azzaytuna University
Tarhuna, Libya

Nasser B. Ekreem
Computer Engineering Department
Azzaytuna University
Tarhuna, Libya

Abstract— in this paper, a new-type recursive least squares algorithm is proposed for identifying the system model parameters and the noise model parameters of Box–Jenkins Systems. The basic idea is based on replacing the unmeasurable variables in the information vectors with their estimates. The proposed algorithm has high computational efficiency because the dimensions of the involved covariance matrices in each subsystem become small. Validation of the model is evaluated using some statistical methods, Which, best-fit criterion and Histogram. Simulation results are presented to illustrate the effectiveness of the proposed algorithm.

Keywords—system identification, model validation, parameter estimation, histogram, Box–Jenkins system

I. INTRODUCTION

Up until the 1950's most of control design relied on Bode, Nyquist and Nichols charts or step response analyses. These methods were mostly limited to SISO systems. In the early 1960's Kalman introduced the state-space representation and through this established state-space based optimal filtering and optimal control theory with Linear Quadratic optimal control as a cornerstone for model-based control design (Gevers, 2006). It was on the heels of this introduction and rise of model-based control that system identification developed [1].

System modeling and system identification are the prerequisite and foundation of all control issues. System identification is about building mathematical models of dynamic systems from observed input–output data. It belongs to the automatic control field with more than 80 years history of theoretical and algorithmic development as well as software packages and industrial applications[2].

System identification contains the structure or order determination and parameter estimation. Parameter estimation is basic for signal filtering, adaptive control and system modelling [3]. System identification has a significant effect on

the filtering, state estimation, system control and optimization. For example, Scarpiniti et al. proposed a nonlinear filtering approach based on spline nonlinear functions; Zhuang et al. presented an algorithm to estimate the parameters and states for linear systems with canonical state-space descriptions; Khan et al. discussed the theoretical implementation of robust attitude estimation for a rigid spacecraft system under measurement loss.

As system identification becomes widely available, many identification methods have been raised, e.g., the gradient identification methods, the hierarchical identification methods, the auxiliary model identification methods and the multi-innovation identification methods[4]. In the field of system identification, the autoregressive (AR) model, moving average (MA) model and autoregressive moving average (ARMA) model are three basic models of the time series.[5]. Multi-stage algorithms have been widely used in the field of identification. For example, Duan et al. presented a two-stage recursive least squares parameter estimation algorithm for output error models, the basic idea is to combine the auxiliary model identification idea and the decomposition technique [6].

The decomposition identification techniques include matrix decomposition and model decomposition can enhance computational efficiencies [7]. Model validation is usually defined to mean “substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model”. A model sometimes becomes accredited through model accreditation. Model accreditation determines if a model satisfies specified model accreditation criteria according to a specified process [8].

During the past few years, Many parameter estimation methods using variety of techniques such as the maximum likelihood methods [9], [10], the least squares methods [11], the gradient methods [12], [13], the subspace identification method [14], [15], the intelligent algorithm [16], [17], the data-

driven identification method [18] and the EM method [19] have been developed [20].

Finally, System identification and parameter estimation methods have been used widely in process control, signal modeling, communication, and electronic technology [21]. The rest of the paper is organized as follows. Section 2, identification model of the proposed algorithm is derived. Section 3, model validation methods is described. In Section 4, simulation results is presented. Finally, concluding remarks are given in Section 5.

II. THE IDENTIFICATION MODEL DERIVATION OF THE PROPOSED ALGORITHM

Consider the following two-stage recursive least squares algorithm for box–Jenkins systems:

$$A(z^{-1})y(k) = B(z^{-1})u(k) + \frac{D(z^{-1})}{C(z^{-1})}n(k) \quad (1)$$

Where

$$A(z^{-1}) = 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_{na}z^{-na}$$

$$B(z^{-1}) = b_1z^{-1} + b_2z^{-2} + \dots + b_{nb}z^{-nb}$$

$$C(z^{-1}) = 1 + c_1z^{-1} + c_2z^{-2} + \dots + c_{nc}z^{-nc}$$

$$D(z^{-1}) = 1 + d_1z^{-1} + d_2z^{-2} + \dots + d_{nd}z^{-nd}$$

are shift operators polynomials is introduced as $z^{-i}y(k) = y(k-i)$, Where $y(k), u(k), n(k)$ are the sequences of system output, measurable input and stochastic input, or noise, respectively, while the constants a_i, b_j, c_i and d_i represent system parameters. The derivation of model identification algorithm represented by eq. (1) is depend on decomposition technique that transform the original identification problem into two sub problems with smaller sizes. First sub problems is system identification model and second sub problems is noise identification model. A System identification model is

$$y_s(k) = -\sum_{i=1}^n a_k y(k-i) + \sum_{i=1}^n b_k u(k-i) \quad (2)$$

This equation can be written as

$$y_s(k) = Z_s^T(k)\theta_s \quad (3)$$

Where

$\theta_s = [a_1, a_2, \dots, a_{na}, b_1, b_2, \dots, b_{nb}]^T$ is vector of system parameters and the information vector is

$$Z_s^T(k) = [-y(k-1) \dots -y(k-na), u(k-1) \dots u(k-nb)]$$

Also, the noise identification model is

$$v(k) = \frac{D(z^{-1})}{C(z^{-1})}n(k) \quad (4)$$

Or

$$v(k) = -\sum_{i=1}^{nc} c_k v(k-i) + \sum_{i=1}^{nd} d_k n(k-i) + n(k) \quad (5)$$

Equation (5) can be written as

$$v(k) = Z_n^T(k)\theta_n + n(k) \quad (6)$$

Where,

$$Z_n^T(k) = [-v(k-1) \dots -v(k-nc), n(k-1) \dots n(k-nd)]$$

And

$$\theta_n = [c_1, c_2, \dots, c_{nc}, d_1, d_2, \dots, d_{nd}]^T$$

By substituting equations (4) and (6) into equation (2), equation (1) can be written as

$$y(k) = Z_s^T\theta_s + Z_n^T(k)\theta_n + n(k) = Z^T\theta + n(k) \quad (7)$$

Where

$$Z^T = \begin{bmatrix} Z_s^T & Z_n^T \end{bmatrix} \quad \text{and} \quad \theta = \begin{bmatrix} \theta_s \\ \theta_n \end{bmatrix}$$

Because the information vector $Z_n(k)$ in $Z^T(k)$ contains unknown variables $v(k-i)$ and unmeasurable noise terms $n(k-i)$. the RLS algorithm cannot be used.

The solution is to replace these unmeasurable variables $v(k-i)$ and $n(k-i)$ in $Z_n(k)$ of $Z^T(k)$ with their estimates as

$$\hat{Z}_n^T(k) = [-\hat{v}(k-1) \dots -\hat{v}(k-nc), \hat{n}(k-1) \dots \hat{n}(k-nd)] \quad (8)$$

$$\text{Then } \hat{Z}^T = \begin{bmatrix} Z_s^T & \hat{Z}_n^T \end{bmatrix}$$

Then, $\hat{v}(k)$ can be computed after system parameters $\hat{\theta}(z)$ be estimated by

$$\hat{v}(k) = y(k) - Z_s^T(k)\hat{\theta}_s(k) \quad (9)$$

The goal is to and to develop the development of new RLS by apply the data filtering technique is required for estimation of parameters. The rational fraction $\frac{C(z)}{D(z)}$ is used and

therefore the model equation (1) becomes ‘‘an equation error model’’, and the recursive least squares algorithm can be applied.

The estimation of $\frac{\hat{C}(z)}{\hat{D}(z)}$ is used to filter the input–output data because $\frac{C(z)}{D(z)}$ is unknown [22]. The identification method is called the filtering based recursive least squares algorithm (F-RLS) [23].

Now, the filtered inputs $u_f(k)$, the filtered output $y_f(k)$ and the filtered $Z_f(k)$ can be defined as

$$u_f(k) = \frac{C(z)}{D(z)}u(k), \quad y_f(k) = \frac{C(z)}{D(z)}y(k) \quad (10)$$

$$Z_f^T(k) = [-y_f(k-1)\dots-y_f(k-na), \\ u_f(k-1)\dots u_f(k-nb)] \quad (11)$$

Equation (2) can be written as

$$A(z)\frac{C(z)}{D(z)}y(k) = B(z)\frac{C(z)}{D(z)}u(k) + n(k) \quad (12)$$

Or

$$A(z)y_f(k) = B(z)u_f(k) + n(k) \quad (13)$$

From this equation, the filtered output is written as

$$y_f(k) = [I - A(z)]y_f(k) + B(z)u_f(k) + n(k) \\ = Z_f^T\theta + n(k) \quad (14)$$

For two identification models equations (14) and equation (6), we can obtain the following recursive least squares algorithm for computing the estimates $\hat{\theta}_s$ and $\hat{\theta}_n$ of θ_s and θ_n :

$$\hat{\theta}_s(k) = \hat{\theta}_s(k-1) + \gamma_f(k)[y_f(k) - Z_f^T(k)\hat{\theta}_s(k-1)] \quad (15)$$

Where

$$\gamma_f(k) = P_f(k)Z_f(k) = \\ [Z_f^T(k)P_f(k-1)Z_f(k) + 1]^{-1}Z_f^T(k)P_f(k-1) \quad (16)$$

And the matrix $P(k)$

$$P_f(k) = [I - \gamma_f(k)Z_f^T(k)]P_f(k-1) \quad (17)$$

With initial conditions

$$P_f(0) = \alpha I \quad \text{and} \quad \theta(0) = 0 \quad (18)$$

With α large ($\alpha = 100, \dots, 1000$)

$$\hat{\theta}_n(k) = \hat{\theta}_n(k-1) + \gamma_n(k)[v_n(k) - Z_n^T(k)\hat{\theta}_n(k-1)] \quad (19)$$

$$\gamma_n(k) = P_n(k)Z_n(k) = \\ [Z_n^T(k)P_n(k-1)Z_n(k) + 1]^{-1}Z_n^T(k)P_n(k-1) \quad (20)$$

$$P_n(k) = [I - \gamma_n(k)Z_n^T(k)]P_n(k-1) \quad (21)$$

The polynomials $C(z)$ and $D(z)$ are unknown, therefore $u_f(k)$, $y_f(k)$, $Z_f(k)$ and $Z_n(k)$ are unknown.

So, the algorithm in equations (15)–(21) cannot be implemented [24]. A solution is to replace the unknown variables with their estimate to derive the new identification algorithms as

$$v(k) = A(z)y(k) - B(z)u(k) = y(k) - Z_s^T(k)\theta_s \quad (22)$$

Substituting equation (6) into equation (22), we get

$$y(k) = Z_s^T(k)\theta_s + v(k) = Z^T(k)\theta + n(k) \quad (23)$$

Replacing θ_s on the right-hand side of equation (22) with its estimate $\hat{\theta}_s(k-1)$, the estimate $v(k)$ can be computed by $\hat{v}(k) = y(k) - Z_s^T(k)\hat{\theta}_s(k-1)$.

Let $\hat{v}(k)$ be the estimate of $v(k)$ and construct the estimate of $Z_n(k)$ as equation (8). From equation (6), we have $n(k) = v(k) - Z_n^T(k)\theta_n$.

Replacing $Z_n(k)$ and θ_n with $\hat{Z}_n(k)$ and $\hat{\theta}_n(k-1)$, the estimate $\hat{v}(k)$ can be computed by $\hat{n}(k) = \hat{v}(k) - \hat{Z}_n^T(k)\hat{\theta}_n(k-1)$. Using the parameter estimates of the noise model

$$\hat{\theta}_n(k) = [\hat{c}_1(k), \hat{c}_2(k), \dots, \hat{c}_{n_c}(k), \hat{d}_1(k), \hat{d}_2(k), \dots, \hat{d}_{n_d}(k)]^T$$

Then, $\hat{u}_f(k) = \frac{\hat{C}(z)}{\hat{D}(z)}u(k)$ and $\hat{y}_f(k) = \frac{\hat{C}(z)}{\hat{D}(z)}y(k)$.

From the previous equations, we can recursively compute $\hat{u}_f(k)$ and $\hat{y}_f(k)$ by the following equations

$$\hat{u}_f(k) = -\hat{d}_1(k)\hat{u}_f(k-1) - \hat{d}_2(k)\hat{u}_f(k-2) - \\ \dots - \hat{d}_{n_d}(k)\hat{u}_f(k-n_d) + u(k) + \hat{c}_1(k)u(k-1) + \\ \hat{c}_2(k)u(k-2) \dots + \hat{c}_{n_c}(k)u(k-n_c). \quad (24)$$

$$\hat{y}_f(k) = -\hat{d}_1(k)\hat{y}_f(k-1) - \hat{d}_2(k)\hat{y}_f(k-2) - \\ \dots - \hat{d}_{n_d}(k)\hat{y}_f(k-n_d) + y(k) + \hat{c}_1(k)y(k-1) + \\ \hat{c}_2(k)y(k-2) \dots + \hat{c}_{n_c}(k)y(k-n_c). \quad (25)$$

Construct the estimate of the $Z_f(k)$

$$Z_f^T(k) = [-\hat{y}_f(k-1), -\hat{y}_f(k-2), \dots, -\hat{y}_f(k-na), \hat{u}_f(k-1), \hat{u}_f(k-2), \dots, \hat{u}_f(k-nb)] \quad (26)$$

Replacing $Z_f(k)$ with $\hat{Z}_f(k)$, and $Z_n(k)$ with $\hat{Z}_n(k)$. Also, $y_f(k)$ with $\hat{y}_f(k)$, and the noise $v(k)$ with $\hat{v}(k)$, two-stage recursive least squares algorithm for box-Jenkins systems is obtained for estimating the parameter vectors θ_s and θ_n for the ARARMA systems:

$$\hat{\theta}_s(k) = \hat{\theta}_s(k-1) + \gamma_f(k)[\hat{y}_f(k) - \hat{Z}_f^T(k)\hat{\theta}_s(k-1)] \quad (27)$$

$$\gamma_f(k) = P_f(k)\hat{Z}_f(k) = [\hat{Z}_f^T(k)P_f(k-1)\hat{Z}_f(k) + 1]^{-1}\hat{Z}_f^T(k)P_f(k-1) \quad (28)$$

In addition, $P_f(k)$

$$P_f(k) = [I - \gamma_f(k)\hat{Z}_f^T(k)]P_f(k-1) \quad (29)$$

$$\hat{\theta}_n(k) = \hat{\theta}_n(k-1) + \gamma_n(k)[\hat{v}_n(k) - \hat{Z}_n^T(k)\hat{\theta}_n(k-1)] \quad (30)$$

$$\gamma_n(k) = P_n(k)\hat{Z}_n(k) = [\hat{Z}_n^T(k)P_n(k-1)\hat{Z}_n(k) + 1]^{-1}\hat{Z}_n^T(k)P_n(k-1) \quad (31)$$

$$P_n(k) = [I - \gamma_n(k)\hat{Z}_n^T(k)]P_n(k-1) \quad (32)$$

$$\hat{v}(k) = y(k) - Z_s^T(k)\hat{\theta}_s(k-1) \quad (33)$$

$$\hat{\theta}_s(k) = [\hat{a}_1(k), \hat{a}_2(k), \dots, \hat{a}_{na}(k), \hat{b}_1(k), \hat{b}_2(k), \dots, \hat{b}_{nb}(k)]^T \quad (34)$$

$$\hat{\theta}_n(k) = [\hat{c}_1(k), \hat{c}_2(k), \dots, \hat{c}_{nc}(k), \hat{d}_1(k), \hat{d}_2(k), \dots, \hat{d}_{nd}(k)]^T \quad (35)$$

The flowchart of computing the parameter estimates $\hat{\theta}_s(k)$ and $\hat{\theta}_n(k)$ is shown in fig. 1. Data from $k=1$ to $k=N$ has been used. Where N is the last value of the samples.

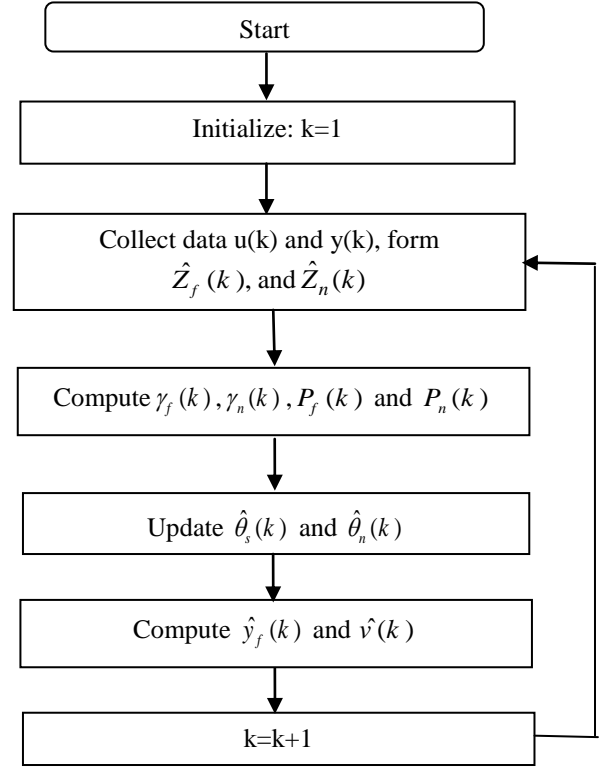


Fig. 1. The flowchart of computing the estimates

III. MODEL VALIDATION

Model validation is the final step of the system identification process; it involves verifying between the measured data and desired data. Based on this work, we will concentrate the study on few criterions [23].

- *Best fit criterion*

The best model structure is the one that minimizes the prediction error. Best fit criterion is often used for model validation, by considering the highest fit [25]. The ideal situation is when the predicted outputs $\hat{y}(k)$ are capable of explaining a major part of the actual output. The ratio R^2 measures the proportion of the actual variation of $y(k)$ that is explained by the regression. It is known as the multiple correlation coefficient and is often expressed in percent [26].

$$R^2 = \frac{\sum_{k=1}^N \hat{y}^2(k)}{\sum_{k=1}^N y^2(k)} = 1 - \frac{\sum_{k=1}^N \varepsilon^2(k)}{\sum_{k=1}^N y^2(k)} \quad (37)$$

Where ε is error between measured and estimated output at the sequence k .

- *Histogram*

Applying statistical pattern recognition techniques often requires an estimation of probability density functions of data samples. There are many methods that have been used for statistical density estimation and one of the common methods is the histogram. The histogram method is the method by which a probability density is constructed from a set of samples, also histogram is easy to create and are computationally feasible [27].

To construct a histogram one needs to select a starting point x_0 and the bin width b and we define the bins of the histogram to be the intervals

$$[x_0 + (i-1)h, x_0 + ih], \quad i = 1, 2, \dots, n.$$

The histogram is defined by:

$$f_{Hist}(x) = \frac{1}{n} \frac{\text{No. of observation in the same bin as } x}{b} \quad (38)$$

More generally, one can also use bins of different widths, in which case

$$f_{Hist}(x) = \frac{1}{n} \frac{\text{No. of observation in the same bin as } x}{\text{Width of bin containing } x} \quad (39)$$

IV. SIMULATION RESULTS

In order to show the performance of the proposed algorithm, consider the following example as a second order system

$$A(z^{-1}) = 1 + 0.25z^{-1} + 0.92z^{-2}$$

$$B(z^{-1}) = -0.9z^{-1} + 0.65z^{-2}$$

$$C(z^{-1}) = 1 + 0.3z^{-1}$$

$$D(z^{-1}) = 1 - 0.45z^{-1}$$

$$\hat{\theta}_s(k) = [0.25, 0.92, -0.9, 0.65]^T$$

$$\hat{\theta}_n(k) = [0.3, -0.45]^T$$

The sequence $u(k)$ is generated as a white sequence with a Gaussian distribution of zero mean and unit variance, while the disturbance $n(k)$ is generated as a white noise sequence with zero mean and variance $\sigma^2 = 0.4$.

The performance of the algorithm is examined when a fault that appear abruptly in a time instant (Abrupt Faults). The parameters of the system will change when the fault happen. In this work, data sequences from $n=0$ to $n=2000$ has been used.

We assume that the fault happen at $n=1000$ in window of the 125 samples. Then, the system parameters are described by

$$\hat{\theta}_s(k) = [0.45, 1.05, -0.5, 0.4]^T$$

$$\hat{\theta}_n(k) = [0.1, -0.25]^T$$

The estimation error is introduced as a measure of the algorithm's effectiveness [2].

where

$$\varepsilon = \sqrt{\frac{\|\hat{a}_1 - a_1\|^2 + \|\hat{a}_2 - a_2\|^2 + \|\hat{b}_1 - b_1\|^2 + \|\hat{b}_2 - b_2\|^2}{\|a_1\|^2 + \|a_2\|^2 + \|b_1\|^2 + \|b_2\|^2}} \quad (40)$$

Fig. 2 shows the estimation errors for this algorithm versus all sequences (k) [28]. The figure shows that the parameter estimation errors for fault free mode and faulty mode. In fault free mode "red color", the parameter estimation errors become smaller and smaller with the sequences increasing, while in faulty mode situation "Black color", the parameter estimation errors become smaller and smaller with the sequences before the occurrence the fault. After the fault occurs, the parameter estimation errors will be increasing with the sequences is increasing.

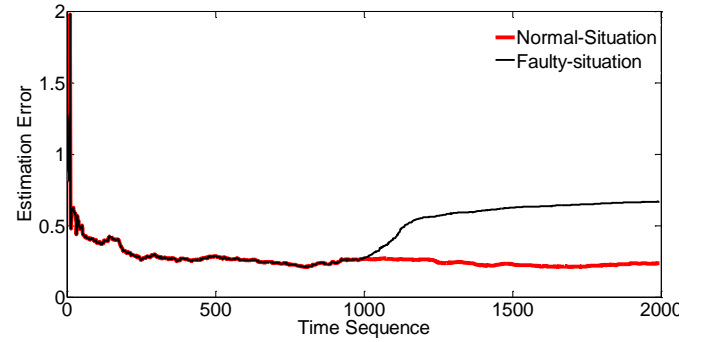


Fig. 2. Estimation errors versus sequences for no fault and faulty situation

The ratio R^2 is computed and the percentage value of best fit for this model equal 92.23%.

Finally, the residual histogram is shown in fig. 3.

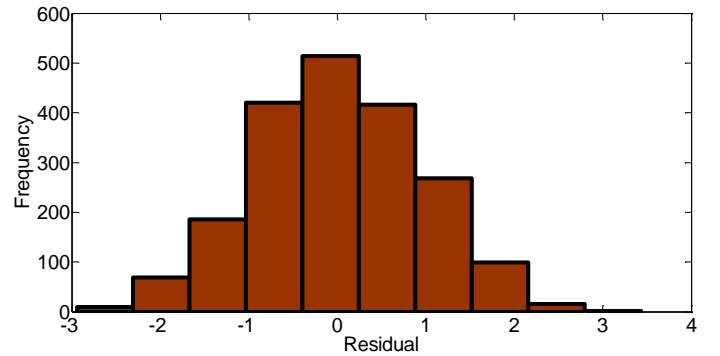


Fig. 3. The residual histogram of the prosed algorithm for validation data

The figure shows that the data is normally distributed, that means, the estimated model describes the observed data well [29].

V. CONCLUSIONS

A new-type recursive least squares algorithm for Box–Jenkins Systems has derived in this paper. The aim is to decompose the system into two subsystems and then interactively to estimate the parameters of each subsystem. This algorithm can improve the computational efficiency due to the input-output data of the system. The simulation results show that the proposed algorithm can generate highly accurate parameter estimates and the application of statistical methods showed that the validation of the model is good. The proposed algorithm can be extended to other linear or nonlinear systems with colored noise.

REFERENCES

- [1] Hernan Guidi, “Open and Closed-loop Model Identification and Validation”, Master Thesis, University of Pretoria December 2008.
- [2] F. Chen and F. Ding, “Recursive Least Squares Identification Algorithm for Multiple-Input Nonlinear Box-Jenkins Systems Using Maximum Likelihood Principle”, *Journal of Computational and Nonlinear Dynamics* Vol. 11, 2016.
- [3] Y. Gu, F. Ding, J. Li, “State filtering and parameter estimation for linear systems with d-step statedelay”, *IET Signal Processing*. Vol. 8, Iss. 6, pp. 639–646, 2016.
- [4] D. Meng and F. Ding, “Model Equivalence-Based Identification Algorithm for Equation-Error Systems with Colored Noise”, *Algorithms*. pp. 280-291, 2015.
- [5] Y. Xiao, Y. Zhang, J. Ding and J. Daic, “The residual based interactive least squares algorithms and simulation studies”, *Computers and Mathematics with Applications*. pp. 1190-1197, 2009.
- [6] S. Wang and R. Ding, “Three-stage recursive least squares parameter estimation for controlled autoregressive autoregressive systems”, *Applied Mathematical Modelling*. pp. 7489-7497, 2013.
- [7] H. Hu, X. Yongsong and R. Ding, “Multi-Innovation Stochastic Gradient Identification Algorithm for Hammerstein Controlled Autoregressive Autoregressive Systems Based on the Key Term Separation Principle and on the Model Decomposition,” *Journal of Applied Mathematics*, Vol. 2013, September 2013.
- [8] R. G. Sargent, “Verification and validation of simulation models”, *Winter Simulation Conference*, January 2011.
- [9] S. Gibson, A. Wills, and B. Ninness, “Maximum-likelihood parameter estimation of bilinear systems,” *IEEE Trans. Autom. Control*, vol. 50, no. 10, pp. 1581_1596, Oct. 2005.
- [10] M. Li, X. Liu, and F. Ding, “The maximum likelihood least squares based iterative estimation algorithm for bilinear systems with autoregressive moving average noise,” *J. Franklin Inst.*, vol. 354, no. 12, pp. 4861_4881, Aug. 2017.
- [11] D. Meng, “Recursive least squares and multi-innovation gradient estimation algorithms for bilinear stochastic systems”, *Circuits Syst. Signal Process.*, vol. 36, no. 3, pp. 1052_1065, Mar. 2017.
- [12] M. Li, X. Liu, and F. Ding, “The gradient-based iterative estimation algorithms for bilinear systems with autoregressive noise”, *Circuits Syst. Signal Process.*, vol. 36, no. 11, pp. 4541_4568, Nov. 2017.
- [13] M. Li, X. Liu, and F. Ding, “Least-squares-based iterative and gradientbased iterative estimation algorithms for bilinear systems”, *Nonlinear Dyn.*, vol. 89, no. 1, pp. 197_211, Jul. 2017.
- [14] D. Westwick and M. Verhaegen, “Identifying MIMO Wiener systems using subspace model identification methods”, *Signal Process.*, vol. 52, no. 2, pp. 235_258, Jul. 1996.
- [15] V. Verdult and M. Verhaegen, “Subspace identification of multivariable linear parameter-varying systems”, *Automatica*, vol. 38, no. 5, pp. 805_814, May 2002.
- [16] Y.-X. Zheng and Y. Liao, “Parameter identification of nonlinear dynamic systems using an improved particle swarm optimization”, *Optik-Int. J. Light Electron Opt.*, vol. 127, no. 19, pp. 7865_7874, Oct. 2016.
- [17] S.-H. Tsai and Y.-W. Chen, “A novel fuzzy identification method based on ant colony optimization algorithm”, *IEEE Access*, vol. 4, pp. 3747_3756, 2016.
- [18] P. T. Brewick and S. F. Masri, “An evaluation of data-driven identification strategies for complex nonlinear dynamic systems”, *Nonlinear Dyn.*, vol. 85, no. 2, pp. 1297_1318, Jul. 2016.
- [19] W. Xiong, X. Yang, L. Ke, and B. Xu, “EM algorithm-based identification of a class of nonlinear Wiener systems with missing output data”, *Nonlinear Dyn.*, vol. 80, nos. 1_2, pp. 329_339, Apr. 2015.
- [20] M. Li, X. Liu, “Auxiliary Model Based Least Squares Iterative Algorithms for Parameter Estimation of Bilinear Systems”, *IEEE Access* Vol. 6, 2018.
- [21] L. Xu, “The parameter estimation algorithms based on the dynamical response measurement data”, *Advances in Mechanical Engineering*. Vol. 9, Iss. 11, pp. 1–12, 2017.
- [22] X. Lu, W. Zhou and W. Shi, “Data Filtering Based Recursive Least Squares Algorithm for Two-Input Single-Output Systems with Moving Average Noises”, *Journal of Applied Mathematics*, vol. 2014, 2014.
- [23] F. Ding and T. Chen, “Parameter estimation of dual-rate stochastic systems by using an output error method,” *IEEE Transactions on Automatic Control*, vol. 50, no. 9, pp. 1436–1441, 2005.
- [24] F. Ding and J. Ding, “Least-squares parameter estimation for systems with irregularly missing data,” *International Journal of Adaptive Control and Signal Processing*, vol. 24, no. 7, pp. 540–553, 2010.
- [25] S. Rachad, B. Nsiri and B.I Bensassi, “System identification of inventory system using ARX and ARMAX models,” *International Journal of control and automation*, December 2015.
- [26] L. Ljung, “System Identification Theory For User”, Prentice Hall Ptr, 1999.
- [27] K. Burnham, D. Anderson, “Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach,” Springer, 2002.
- [28] G. Yoo and R. Ding, “Two-stage least squares based iterative identification algorithm for controlled autoregressive moving average (CARMA) systems”, *Computers and Mathematics with Applications*. Vol. 8, Iss. 6, pp. 975–984, 2012.
- [29] H. Shariff, M. Rahiman, I. Yassin And M. Tajjudin, “System Identification Of A Steam Distillation Pilotscale Using Arx And Narx Approaches,” *International Journal Of, Research And Technology*, Vol. 03, Issue:01, 2014.