ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

NONLINEAR IDENTIFICATION OF A BATCH REFLUX RATIO CONTROLLED DISTILLATION COLUMN

Ilham Rustam, Wan Ahmad Khusairi Wan Chek, Mohd Saiful Najib Ismail Marzuki, Mohd Aldrin Ali and Mohd Nazrul Sidek
Faculty of Electrical Engineering, Universiti Teknologi Mara, Dungun, Terengganu, Malaysia
E-Mail: drilhamrustam@gmail.com

ABSTRACT

An attempt towards representation of a nonlinear process dynamic through the use of empirical NARX-OLS approach is discussed in this paper. Structured data mining was incipiently conducted to provide exposure over the process properties elucidated by the plant taken from a MISO control perspectives that is allocated in accordance to various sets of reflux ratio operation. Establishment over the process dynamic identification is made possible via OLS model structure and QR factorization parameter selection technique. An optimization of the resultant finding was then implemented through ERR optimization for comparison study. A good comparable yielded result from the model estimators insinuated the possibility of constructing a well-defined empirical dynamic model through a structured data mining processes without a priori knowledge of the system.

Keywords: nonlinear system identification, distilation column, NARX model, OLS-ERR model.

INTRODUCTION

Dynamic modelling is a field of study that engages in building a mirroring image of a process dynamic through mathematical approach that is based upon the underlying input and output relationships [1]. An empirical model is a classification of system identification that depends on prior knowledge of the system which can be divided into white, grey and black box modelling[2].

There exist various approaches reported over the years that have successfully implemented system identification to capture a given process dynamics. Among the most adopted is the Auto Regressive with eXogenous inputs (ARX), Nonlinear Auto Regressive Moving Average with eXogenous inputs (NARMAX), Nonlinear Auto Regressive with eXogenous inputs (NARX) and Nonlinear Auto Regressive with Moving Average (NARMA). Extensive literatures are available that addresses these approaches application found in various engineering discipline, economics and applied sciences studies owing to each powerful and efficient representation of a system [3-8].

This paper discusses a forayinto an emulation of a water-ethanol distillation process pilot plant using a well-established nonlinear identification technique to ultimately achieve regularization of the end top product purity. An adaptation of the Nonlinear Autoregressive models with eXogenous input (NARX) model approach is herewith implied, which will be based upon the Orthogonal Least Squares (OLS) technique to determine the model structure selection guided by criterion discussed in [9].

This paper is organized as follows: background of research presented in section 2, followed by methodology in section 3 and result with discussion in section 4. Finally, the concluding remarks in section 5.

BACKGROUND OF RESEARCH

The NARX adoption in this study due to its ability to conveniently describe a nonlinear system as reported in numerous implementation records [10-11]. It is a subset of Nonlinear Autoregressive Moving Average with Exogenous Inputs (NARMAX) [20] modelling, differ such that the incorporated related noise model was reduced into a single term, $e(t)$. Great flexibility of it polynomial forms can be imagined in a nonlinear time-invariant system as

$$y(t) = f^{d}\begin{bmatrix} y(t-1), \dots y(t-n_y), \\ u(t-1), \dots u(t-n_u) \end{bmatrix} + \varepsilon(t) \quad (1)$$

$$\begin{aligned}\varphi(t) &= [y(t-1), \dots y(t-n_y), u(t-1), \dots u(t-n_u)]^T \\ &= [x_1(t), \dots x_{n_y}(t), x_{n_{y+1}}(t), \dots x_n(t)]^T \end{aligned} \quad (2)$$

where the delayed output-input data are $y(t-1), .. y(t-n_y), u(t-1), . u(t-n_u)$, made essential using the input, $u(t)$, output, $y(t)$ and white noise residuals, $\varepsilon(t)$. The function f^d is the unknown nonlinear coefficient and $\varphi(t)$ is the model regression vector with n_y and n_u as the associated lag of output and input terms respectively. Interest on undertaking the proposed approach on nonlinear application modelling can be derived from literatures such as [12-13], discussing on variance decomposition through selective exogenous and endogenous model candidate terms.

As the degree of multivariate polynomial increases following the highest order among the terms, indicated by l, where there are marginal incremental in n-dimension of regression vector $\varphi(t)$, a redundancy or insignificant contribution of the candidate model term could be identified with several literatures [14-15] cited in favour of eliminating said term from the eligible candidate model terms.

METHODOLOGY

Hardware description

VOL. 12, NO. 10, MAY 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

A distillation pilot plant, model number BP 681-50, developed primarily for teaching, training and research purposes was used to substantiate the findings of this study. It is engaged in mixture separation at atmospheric pressure in both batch and simulated continuous operation through the use of the reboiler and feed product tank, B1 and B2. In this study however, only the batch operation is observed and is achieve via two electrical cartridge heaters that supplies direct heat input to the batch feed mixture in reboiler unit, B1. Upon reaching the boiling point of the more volatile component, vapours would form and would subsequently rise across the tower. Further separation efficiency is achieved at each stages of the bubble cap tray from TT110 (Tray 10) to TT112 (Tray 12) along the tower until the vapour reaches the condenser unit W2 at the top section. At this stage, the hot vapour would be cooled via condensation with cooling water flow, converting it into a cooled liquid form which then would be return back to the tower as a down pouring liquid. This would further help purify the separation process of the end product, as the process is being repeated which in practice is referred to as total reflux operation.

Typically this state of operation is kept continuous for a period of time until such the reflux divider valve KFS-101 has been signalled to open allowing the distillated product to escape the cycle through its orifice and be accumulated in the distillate tank, B4. This event signifies a distillate operation and the whole process can be conducted periodically to assume a cyclic operation or varying reflux ratio [16-17] in which the latter is adapted into this study. The separation behaviour can be monitored through available sensors allocated across the column consisting of each tray stages temperatures, distillate and bottom product flow (Fl-302 and Fl-303) as well as tower pressure sensor (dPT-201). These parameters provided an accommodation for real time monitoring of each variable and are therefore used as inputs to model the process under study.

The implementation of the system identification works are done on an Intel® Core ™ i7-4700MQ CPU with processing power of 2.4GHz, 8Gb RAM laptop computer equipped with a pre-installed 64-bit Win 8.1 operating software and MATLAB 2013 revision A.

Dataset description

The modelling input datasets were obtained experimentally from the distillation pilot plant in an effort to control the top product purity of a water-ethanol mixture via an indirect control approach in a multiple input single output (MISO) control assumptions. This study put its focus primarily on reflux ratio control perturbations as opposed to other approaches. This is done to observe optimal operation time given pre-determined amount of batch binary mixture as well as to avoid the implications cited by [18] whom discussed on the risk of gaining heavier key component inside the distillate product resulted from a rise in top section temperature during a particular process run. However, the reflux ratio implementation is physically limited by the factory installed timer-based digital controller as opposed to plant setup that is widely employed in [16-17]. It is mainly due to the unconventional distillate adaptation at the top section of the plant and the inexistence of mechanisms to measure the product flow rendering conventional numerical calculation of internal reflux ratio to be unapproachable.

The sampling time for each real time monitoring of process variables was set at 5 seconds interval, recorded digitally using the SOLDAS® software. 5 minutes interval was allocated between each process disturbances with each perturbation kept at constant for 4 iterations. Each of the sets would then be translated into a 0 to 5 volt DC signals that is transmitted to the reflux divider valve to essentially indicate an open (distillate) or a close (reflux) operation appropriately. As can be seen in Figure-1, it is apparent that the extension of perturbation signals is mainly to achieve manipulation over different sets of durations required for each states of operation. The set reflux ratio perturbations is executed continuously until it is clear that the light key component in the batch feed mixture has been exhausted.

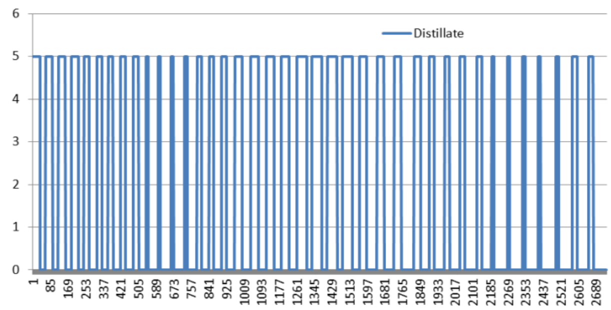
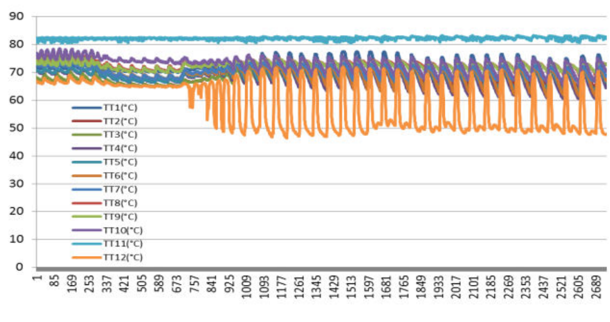


Figure-1. 0 to 5 volt DC signal transmitted to reflux valve.

The plotted data in Figure-2 elucidate each stages of temperature variables during, which the process perturbation was introduced along with the pressure variable and the resultant top product purity defined in RI values respectively. The datasets were partitioned using interweaving technique, to yield two separate entities of data, which from here onwards shall be referred to as the training and validation dataset as part of the pre-processing methods employed in this study.



ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

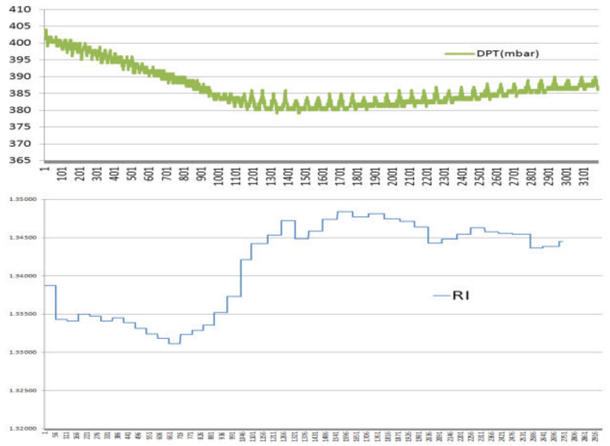


Figure-2. Temperature, pressure and top product purity data plot.

Model selection

It is common knowledge that in real practices, there could be minimal to no available pre-ordained process mechanic insights for a given identification study to take advantage of, as such is the current process of interest. This presented an apparent lack of abbreviations in lagged input and output data terms (n_u,n_y) to essentially placate the methods requirement due to the unavailability of a priori knowledge. Thus, it is justifiable for this study to consider an iterative solving method conducted in finite solving sequences to arrive at the optimal model solution, sorted successively through each iterations calculated MSE value. The best model performance would then be selected from each resultant estimated model permutations that yielded the lowest testing set MSE score value represented as:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad (3)$$

which measures the average squares of errors between the estimated and its actual value [19].

Model validation

Quantification of performance differences between implied approaches in this work are analyse using various well establish residual analysis techniques, of which among others are visualized through the use of goodness of fit test between the output of trained estimated model and the actual output data. The higher percentage of output variance achieved translates to a better model performance. Further validation was done using Akaike's Final Prediction Error (FPE) to compare both linear and nonlinear models. In theory, the smaller the resultant value, the better and more accurate the model. This work also employs minimization of Akaike's Information Criterion (AIC) and Rissanen's Minimum Description Length (MDL), along with Auto- correlation function (ACF) and cross-correlation test (CCF) as model validation tools to ensure the estimated model are comprehensively define and the residuals has no correlation with the model parameters, therefore satisfying whiteness test and independence test criteria.

RESULTS AND DISCUSSIONS

This section investigates the performance of the adopted NARX approach to comprehend its ability to represent a multivariable nonlinear system dynamics. Utilizing iterative model parameter solving approach to produce a pool of solved model estimators for model order of up to 100, the resultant optimal model structure was found at 38 for lagged output terms and 16 for lagged input terms from a total of 262 candidate model terms as guided by the least amount of MSE value.

Table-1. AIC, FPE, MDL, MSE and best fit results of testing and training dataset.

	AIC	**FPE**	**MDL**	**MSE**	**OSA**
Training Set	0.6037	0.6953	0.6406	0.8817	98.7696
	AIC	FPE	MDL	MSE	OSA
Testing Set	0.8	0.9213	0.8489	1.1684	98.371

Table-1 presented a consignment of the estimated model performance result through the use of AIC, FPE, MDL, MSE and R-squared validation measures. It can be seen that the modelling technique produces a preferable small resultant value across the test prompted in this study, with R-squared test indicated a high percentage of accuracy substantiated by more than 98% for both the training and testing datasets. Although testing dataset has scored less in overall than its counterpart, these results are deemed promising given that the datasets is considered as an unforeseen perturbation process and the resultant value yielded is considered as a minor discrepancy of otherwise an acceptably good small figures therefore provided an indication that the model has been able to mitigate a realistic offline representation of the plant. Figure-3 would then qualitatively illustrate the one-step-ahead predictions of the developed model and again as has been mentioned, the resultant model has manage to follow the actual data plots patterns given some margin of errors, in both scenarios without vital apparent differences between the training and testing prediction plot.

It can also be seen in Figure-4 that the residuals of the approach has shown an acceptable performance based on residual auto-correlation analysis performed albeit an out of bound on lag terms 1 and 7 of the testing set and on lag terms 12 for both cross-correlation test. A review on the histogram analysis in Figure-5 indicates that a slight skew to the left of otherwise a Gaussian shaped result has been obtained by the estimated model. This insinuated that the residuals of the model is a white noise in nature, and therefore prove that the model has properly encapsulated the underlying process dynamics.

VOL. 12, NO. 10, MAY 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

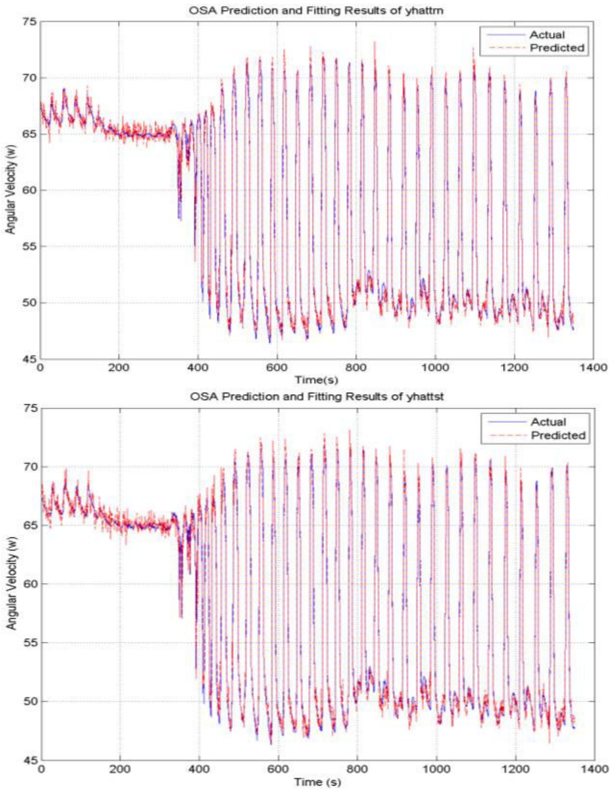


Figure-3. OSA results of training and testing dataset from *NARX* approach.

CONCLUSIONS

The identification of water-ethanol separation process dynamics in BP 681-50 distillation column pilot plant has successfully been mitigated in this work. The case study presented a complex identification problem due to the established nonlinear relationships and amount of variables involved that afforded this study to partake in a credible case of nonlinear multivariable identification process. The yielded results had also imparted a reflection on the success of data mining structure that was executed in divulging the plant dynamic properties which can be seen through the model validation measures performed on the estimated model. Ultimately, it is proven that the NARX-OLS approach is capable to mitigate practical application of system identification where there is no a priori system structure knowledge existed.

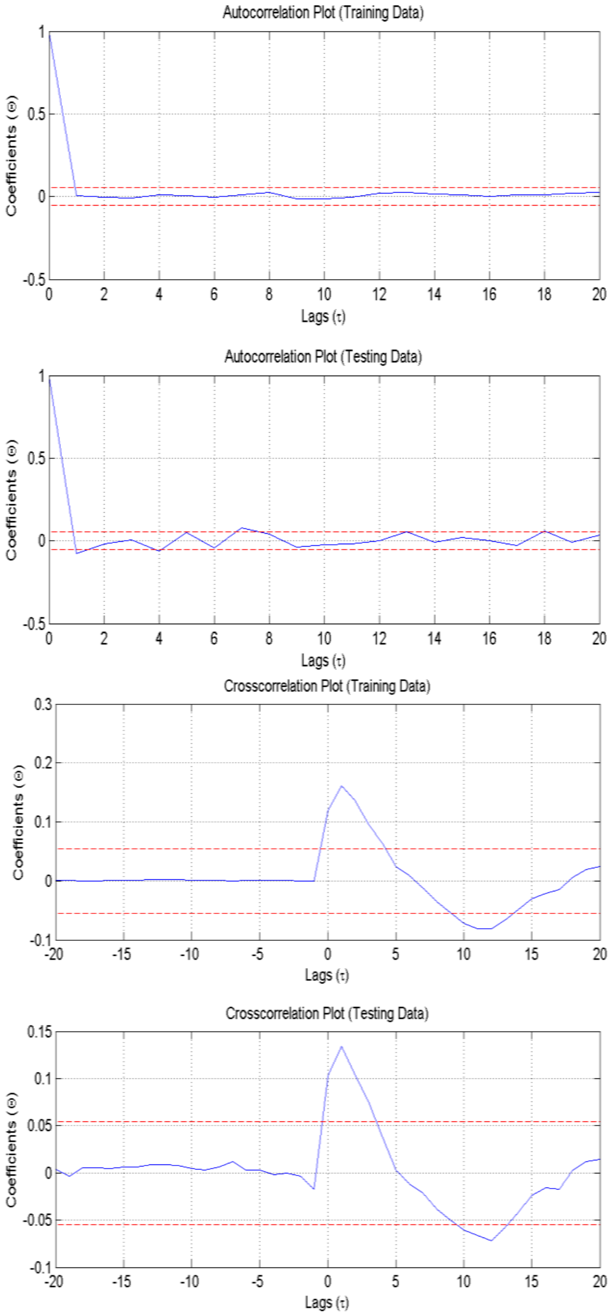


Figure-4. Autocorrelation and cross correlation results of testing and training dataset from *NARX* approach.

ARPN Journal of Engineering and Applied Sciences

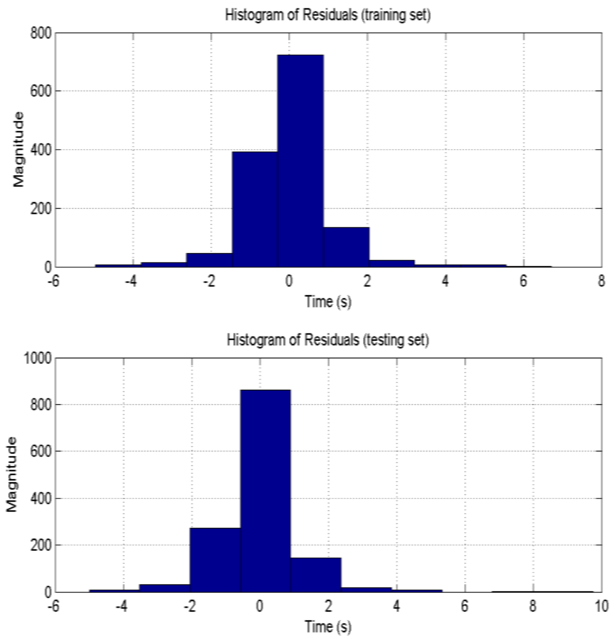


Figure-5. Histogram analysis results of testing and training dataset from *NARX* approach.

ACKNOWLEDGEMENT

The author would like to acknowledge the financial support of Research Acculturation Grant Scheme (RAGS) number: 600-RMI/RAGS 5/3/(205/2013) from Research Management Institute (RMI) of UiTM, Malaysia.

REFERENCES

[1] Ljung L. 1999. System identification: Theory for the user. 2nd Ed. Prentice-Hall, New Jersey, USA.

[2] K. Worden, C. X. Wong, U. Parlitz, A. Hornstein, D. Engster, T. Tjahjowidodo, F. Al-Benderc, D. D. Rizosd and S. D. Fassois. 2007. Identification of pre-sliding and sliding friction dynamics: Grey box and black-box models. Mechanical Systems and Signal Processing. 21(1): 514-534.

[3] Zheng S., Zheng W. and Jin X.2007. An application of system identification method on a regional economic system. In IEEE International Conference on Wireless Communication, Networking and Mobile Computing. pp. 3875-3878.

[4] A. Giwa and S. Karacan. 2012. Nonlinear black-box modeling of a reactive distillation process. International Journal of Engineering Research and Technology.1 (7): 1-10.

[5] S. R. Anderson, N. F. Lepora, J. Porrill and P. Dean. 2010. Nonlinear dynamic modeling of isometric force production in primate eye muscle. IEEE Transactions on Biomedical Engineering.57 (7): 1554-1567.

[6] M. A. Balikhin, R. J. Boynton, S. N. Walker, J. E. Borovsky, S. A. Billings and H. L. Wei. 2011. Using the NARMAX approach to model the evolution of energetic electrons fluxes at geostationary orbit. Geophysical Research Letters. 38(18): 1-5.

[7] B. H. G. Barbosa, L. A. Aguirre, C. B. Martinez and A. P. Braga. 2011. Black and gray-box identification of a hydraulic pumping system. IEEE Transactions on Control Systems Technology. 19(2): 398-406.

[8] J. O. Hahn, D. B. McCombie, A. T. Reisner, H. M. Hojman and H. H. Asada. 2010. Identification of multichannel cardiovascular dynamics using dual laguerre basis functions for noninvasive cardiovascular monitoring. IEEE Transactions on Automatic Control Technology. 18(1): 170-176.

[9] Taib M. N. 1993. Time series modelling and prediction using neural networks. Master thesis, University of Sheffield, University of Sheffield, England.

[10] Yassin I. M., Taib M. N., Rahim N. A., Salleh M. K. M. and Abidin H. Z. 2010. Particle swarm optimization for NARX structure selection-Application on DC motor model. In: IEEE Symposium on Industrial Electronics and Applications. pp. 456-462.

[11] Yassin I. M., Taib M. N., Hassan H. A., Zabidi A. and Tahir N. M. 2010. Heat exchanger modeling using NARX model with binary PSO-based structure selection method. In: International Conference on Computer Applications and Industrial Electronics. pp. 368-373.

[12] Wei H. L. and Billings S. A. 2008. Model structure selection using an integrated forward orthogonal search algorithm assisted by squared correlation and mutual information. International Journal of Modelling, Identification and Control. 3(4): 341-356.

[13] T.J. Harris and W. Yu. 2012. Variance decompositios of nonlinear time series using stochastic simulation and sensitivity analysis. Statistics and Computing. 22(2): 387-396.

[14] Billings S. A. and Wei H. L. 2008. An adaptive orthogonal search algorithm for model subset

selection and non-linear system identification. International Journal of Control. 81(5): 714-724.

[15] Cheng Y., Wang L. and Hu J. 2009. A two-step method for nonlinear polynomial model identification based on evolutionary optimization. In: World Congress on Nature and Biologically Inspired Computing. pp. 613-618.

[16] Perry R. H. and Green D. W. 2007. Perry's chemical engineer' handbook. 7th Ed. McGraw-Hill, New York, USA.

[17] Sørensen E. and Skogestad S. 1994. Optimal operating policies of batch distillation with emphasis on the cyclic operating policy. In: 5th International symposium on Process Systems Engineering. pp. 449-456.

[18] Kister H. Z.1990. Distillation operation. McGraw-Hill, New York, USA.

[19] Wackerly D. and William S. 2008. Mathematical statistics with applications. 7th Ed. Thomson Higher Education, California, USA.

[20] I. M. Yassin, A. Zabidi, M. S. A. M. Ali, N. M. Tahir, H. A. Hassan, H. Z. Abidin and Z. I. Rizman. 2016. Binary particle swarm optimization structure selection of nonlinear autoregressive moving average with exogenous inputs (NARMAX) model of a flexible robot arm. International Journal on Advanced Science, Engineering and Information Technology. 6(5): 630-637.