International Journal of Chemical Sciences and Applications
ISSN 0976-2590, Online ISSN 2278 – 6015
Vol 6, Issue2, 2015, pp 64-71
http://www.bipublication.com

Research Article

OLS based-RBF Networks for Estimation of Product Composition in Batch Distillation

Nileshkumar. N. Yelgatte[a, b] Sunil. A. Raut[b],
Shubham Mehta[a] and Imran Rahman*[a]

[a]Chemical Engineering and Process Development Division,
CSIR-National Chemical Laboratory, Dr. Homi Bhaba Road, Pune-411008, India
[b]University Institute of Chemical Technology, N. M. U. Jalgaon-425001, India
*Corresponding author: Email: r.imran@ncl.res.in, Tel.: 91-020-25902428; Fax: 91-020-25902675

[Received-11/06/2015, Accepted-25/06/2015, Published-07/08/2015]

ABSTRACT:

Orthogonal Least Square (OLS) based Radial Basis Function (RBF) network is developed for a batch distillation process to estimate the product compositions from temperature measurements. The dynamics of the batch distillation was generated by varying the initial feed composition and boilup rate from batch to batch and within each batch reflux ratio was also randomly changed. The results show the potential of the method for developing a soft sensor for chemical processes.

Keywords: Batch Distillation; Soft Sensor; Composition Estimation; Radial Basis Function; Orthogonal Least Square regression

1. INTRODUCTION

Batch distillation is an important unit operation that is widely used in fine and specialty chemical, pharmaceutical industries, where product volume is small and strict purity is required. Thus, it is necessary to monitor the composition in distillate stream during the entire batch operation. This is difficult due to the lack of quick and consistent online composition analyzers, so it is better to use online state estimation techniques to provide reliable and real-time purity estimates from readily available temperature measurement.

Several papers have addressed the issue of state composition estimation of products in a batch distillation process. Quintero-Marmol, Luyben, and Georgakis (1991) and Quintero-Marmol and Luyben (1992) used a quasi- dynamic estimator (QDE) and an extended Luenberger observer (ELO) to estimate the distillate composition for a ternary batch column by using first principle model [14-15]. In order to improve the estimators robustness to process/model mismatch and measurement noise Barolo, Pistillo, and Trotta (2000), Oisiovici and Cruz (2000, 2001) and Venkateswarlu and Avantika (2001), developed an extended Kalman filter (EKF) to estimate the product composition by using temperature

Cite this article- Nileshkumar. N. Yelgatte, Sunil. A. Raut, Shubham Mehta and Imran Rahman, OLS based-RBF Networks for Estimation of Product Composition in Batch Distillation, 2015, International Journal of Chemical Sciences and Applications, Vol 6, Issue2, 2015, pp 64-71 http://www.bipublication.com

measurement [16]. EKFs are much more robust to mismatch and noise, but still their performance heavily depends on the thermodynamic modeling of Vapour Liquid Equilibrium (VLE) [12-13]; they are also difficult to initialize and tune, and require considerable computational effort for on-line use (E.Zamprogna, Barlo and Seborg, 2004) [1-8].

E.Zamprogna et al. (2004), studied various Partial Least Squares (PLS) approach as an alternative to first-principle model, to estimate the quality of products obtained from a batch distillation column not only at the end of the process, but also during the entire duration of the batch. E. Zamprogna, Barlo and Seborg (2004) also used separate PLS regression algorithms for different portions of the batch duration i.e. startup, main cut and slop cut, but a major drawback of PLS regression is that it makes poor predictions when relationship between process input and output is non-linear [8].

Neural Networks have been used in many areas of chemical engineering. It is popular, as it can realize some complex non-linear decision function and approximate certain complicated data-generating mechanisms. A disadvantage of neural networks is that they are highly nonlinear in the parameters. A viable alternative to highly non-linear-in-the parameter neural networks is the Radial Basis Function (RBF) network. RBF method had traditionally been used for strict interpolation in multidimensional space [7].

The performance of RBF network critically depends upon the chosen centers. In, practice, the centers are often chosen to be a subset of the data (Chen et al., 1990) [4]. A learning strategy based on the Orthogonal Least Square (OLS) algorithm provides a systematic approach to the selection of RBF centers [5].

This method selects regressors in a forward manner by the maximization of model error reduction ratio. It has been shown that the learning strategy offers a powerful procedure for fitting adequate and parsimonious Radial Basis Function (RBF) networks (Chen, Billings & Luo, 1989; Billings,Chen, & Korenberg, 1989; ; Chen, Cowan & Grant 1991) [2-3].

The OLS algorithm is a structural identification algorithm and it constructs an adequate network structure in an intelligent way during learning and overcomes the initialization problem.

It shows the contribution of an individually selected model regressor to the modeling accuracy and is an efficient implementation of forward regression procedure for subset model selection. The aim of this work is to develop a multi-input multi-output OLS based-RBF model applicable to entire operation of batch distillation process to predict instantaneous mole fraction of the light (x_{D1}) and intermediate components in the distillate stream (x_{D2}), and mole fraction of the heavy component (x_{B3}) in the reboiler. It is supposed that temperature measurements are available from the reboiler and three trays. The data set need to develop the model were generated by first-principles model of the batch column by varying initial feed composition and boilup rate from batch to batch and reflux ratio within the batch, which will be more representative of batch distillation process. In this paper, we use OLS based RBF networks to estimate the quality of the products as a function of selected trays and reboiler temperature measurements.

The paper is organized as follows: Section 2 provides a brief review to Orthogonal Least Square model. Section 3 describes process description. Section 4 explains the data generation. Section 5 gives the simulating results and discussions. The section 6 concludes the paper.

2. Orthogonal Least Square model

The multi-output RBF neural network model can be formulated as (Billings, Chen & Korenberg, 1989) [3-6].

$$y_i(k) = \hat{y}_i(k) + e_i(k) = \sum_{j=1}^{M} \theta_{j,i}\phi_j(k) + e_i(k)$$

$$1 \le k \le N \qquad (1)$$

for $1 \le i \le n_o$, where $y_i(k)$ is the ith target or desired output, is the error $e_i(k)$ is the error between $y_i(k)$ and the ith model output $\hat{y}_i(k)$, are the model weights $\theta_{j,i}$, $\phi_j(k)$ are the regressors, M is the total number of candidate regressors, n_o the number of outputs, and N the number of training samples. Define

$$\mathbf{y}_i = \begin{bmatrix} y_i(1) \\ y_i(2) \\ \vdots \\ y_i(N) \end{bmatrix}, \qquad \mathbf{e}_i = \begin{bmatrix} e_i(1) \\ e_i(2) \\ \vdots \\ e_i(N) \end{bmatrix},$$

$$\boldsymbol{\theta}_i = \begin{bmatrix} \theta_{1,i} \\ \theta_{2,i} \\ \vdots \\ \theta_{M,i} \end{bmatrix} \qquad (2)$$

for $1 \le i \le n_o$

$$\boldsymbol{\Phi} = [\Phi_1 \;\; \Phi_2 \;\; \cdots \Phi_M]$$

with

$$\Phi_j = [\phi_j(1) \;\; \phi_j(2) \;\cdots \phi_j(N)]^T \qquad (3)$$

$$1 \ge j \ge M \qquad (4)$$

The multi-output regression model (1) becomes

$$\mathbf{y}_i = \boldsymbol{\Phi}\boldsymbol{\theta}_i + \mathbf{e}_i, 1 \le j \le n_o \qquad (5)$$

Further define

$$\mathbf{Y} = [\mathbf{y}_1 \;\; \mathbf{y}_2 \;\; \cdots \mathbf{y}_{no}], \qquad \boldsymbol{\Theta} = [\boldsymbol{\theta}_1 \;\; \boldsymbol{\theta}_2 \;\; \cdots \boldsymbol{\theta}_{no}],$$

$$\mathbf{E} = [\mathbf{e}_1 \;\; \mathbf{e}_2 \;\; \cdots \mathbf{e}_{no}]. \qquad (6)$$

The regression model (1) can be rewritten as

$$\mathbf{Y} = \boldsymbol{\Phi}\,\boldsymbol{\Theta} + \mathbf{E} \qquad (7)$$

Let an orthogonal decomposition of the regression matrix $\boldsymbol{\Phi}$ be

$$\boldsymbol{\Phi} = \mathbf{W}\mathbf{A} \qquad (8)$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & a_{1,2} & \cdots & a_{1,M} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{M-1,M} \\ 0 & \cdots & 0 & 1 \end{bmatrix} \qquad (9)$$

and

$$\mathbf{W} = [\mathbf{w}_1 \;\; \mathbf{w}_2 \;\; \cdots \mathbf{w}_M] \qquad (10)$$

with orthogonal columns that satisfy $\mathbf{w}_j^T\mathbf{w}_l = 0$, if $j \ne l$. The regression model (7) can alternatively be expressed as

$$\mathbf{Y} = \mathbf{W}\mathbf{G} + \mathbf{E} \qquad (11)$$

where the orthogonal weight matrix

$$\mathbf{G} = [\mathbf{g}_1 \;\; \mathbf{g}_2 \;\; \cdots \mathbf{g}_{no}] \qquad (12)$$

with

$$\mathbf{g}_i = [g_{1,i} \;\; g_{2,i} \;\; \cdots g_{M,i}]^T, 1 \le i \le n_o \qquad (13)$$

and $\mathbf{G}$ satisfy the triangular system [6],

$$\mathbf{A}\,\boldsymbol{\Theta} = \mathbf{G} \qquad (14)$$

Knowing $\mathbf{A}$ _ and $\mathbf{G}$, $\boldsymbol{\Theta}$ _ can readily be solved from (14)

To determine extent of correlation between input and output data by the OLS, the root mean squared error (RMSE) is defined as,

$$\text{RMSE} = \sqrt{\frac{\sum_{k=1}^{N}(y_k - y_p)^2}{N}} \qquad (15)$$

Where y_k and y_p are desired and predicted value respectively, for the k[th] output variable and its estimate obtained from OLS model, respectively and *N* is the number of input-output patterns.

3. Process Description

The present work describes the separation of a ternary mixture of *n*-heptane, cyclohexane and toluene by using batch distillation. The process consists of a reboiler, a batch rectifier with 17 trays, a total condenser and four distillate accumulators (Fig.1). The process is operated as per the constant reflux strategy described by Luyben (1991) [10]. The column is initially operated at total reflux when the distillate stream meets the desired purity specification, the distillate withdrawal is started; products and slope cuts are sequentially collected from the top and segregated in distillate accumulators. The heaviest cut is recovered from the reboiler at the end of the batch. Thus, cyclohexane (light product) and *n*-heptane (intermediate product) are accumulated in the two main cut accumulators, respectively, during the batch. Toluene (heavy product) is received from the reboiler residue at the end of the batch.

The batch distillation process is modelled by a set of Differential Algebraic Equations (DAE). Theoretical trays, equimolal overflow are assumed and UNIQUAC thermodynamic model is used to correlate the Vapour Liquid Equilibrium (VLE). The UNIQUAC interaction parameter was calculated from binary parameters reported by Jimenez, Basualdo, Gomez, Toselli and Rosa (2002) [9]. The minimum product purity specifications are 89.5, 86.3 and 99 mol% of cyclohexane, *n*-heptane and toluene, respectively. The batch size of 2930 mol with feed composition of cyclohexane, *n*-heptane and toluene, tray holdup of 7.5 mole, reflux drum holdup of 44 mole with boilup rate of 2750 mole is considered for the present study. The process configuration is similar to a study described by Mujtaba and Macchietto (1996) [11]

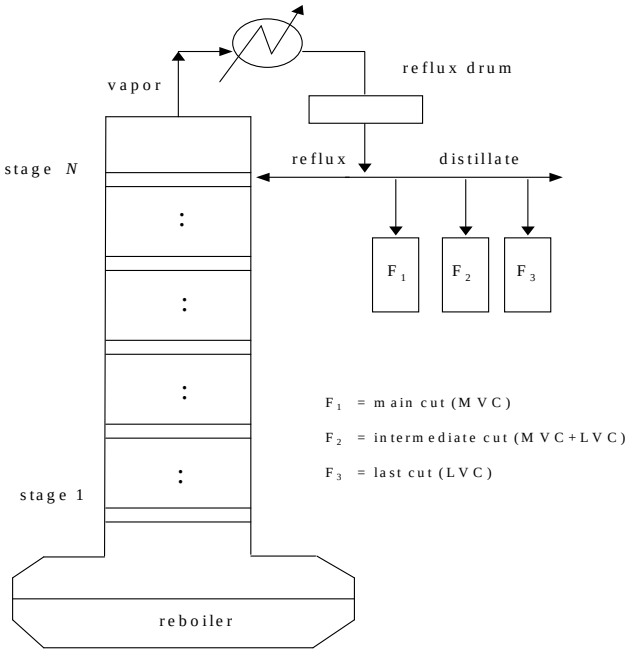


Fig. 1. Schematic of Batch Distillation.

4. Data Generation

The data sets are needed to develop the correlation between purity of the distillate and tray temperatures by OLS regression. The input-output data was generated by running the first-principles model (Luyben, 1991) [10] under different operating conditions, which consists of a set of ODEs. The stiff-ODE, implicit numerical-integration-algorithm LSODE is used to integrate the ODEs.

Nine batch operations were simulated by changing the initial feed composition (x_f) and boilup rate from batch to batch as shown in Table 1. For each batch run, the time-varying data were recorded using sample times of 0.25 h, changing reflux ratio (r) arbitrarily with a switching probability of $P_s = 0.90$, with the values at each transition drawn from a uniform distribution in the range $r = [0.85, 0.99]$. The reflux ratio is defined as ratio of liquid refluxed and vapor rate (L_R/V).

The input-output data were partitioned into two sets, i.e. training and test sets. The training and

test set consisted of five and four batches, respectively.

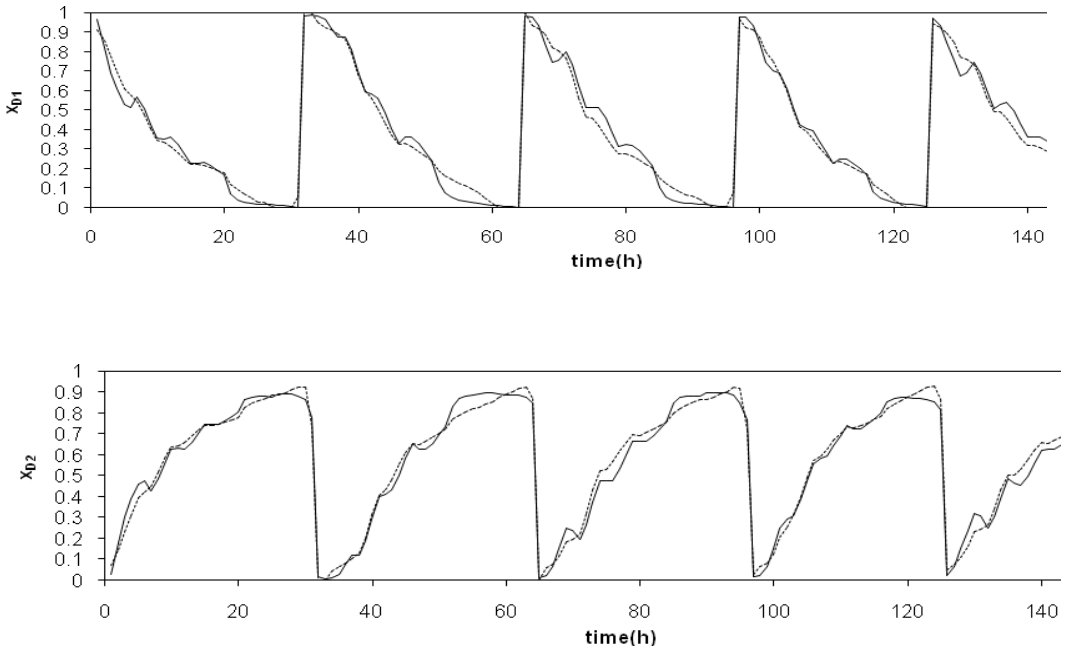
Thus, a highly informative process database was generated easily and economically consisting of temperature at bottom (T_B), temperature at 5th (T_5), 10th (T_{10}), 17th plate (T_{17}) as inputs; mole fraction of cyclohexane and n-heptane in distillate (x_{D1}, x_{D2}) and toluene in reboiler (x_{B3}) as outputs.

Table 1 Data set used for batch distillation simulation

Batch	x_f	V **(kmol/h)**
Training data		
TR01	0.30/0.50/0.20	2.80
TR02	0.43/0.42/0.15	2.63
TR03	0.40/0.42/0.18	2.83
TR04	0.37/0.44/0.19	2.74
TR05	0.36/0.48/0.16	2.64
Testing data		
TS01	0.39/0.40/0.21	2.78
TS 02	0.34/0.45/0.21	2.80
TS 03	0.43/0.40/0.17	2.76
TS 04	0.39/0.41/0.20	2.73

5. RESULTS and DISCUSSION

The OLS based RBF network has been utilized to correlate the reboiler and tray temperatures and mole fraction of the product in a batch distillation process. In batch distillation column, the process dynamic regime moves from a condition of total reflux to a condition of constant reflux. At the same time, the process experiences large excursions in the tray compositions, due to the movement of the light and intermediate components from the bottom to the top of the column. These changes in the column dynamic regime and composition distribution are reflected in the tray temperatures. These tray temperatures are taken as input to



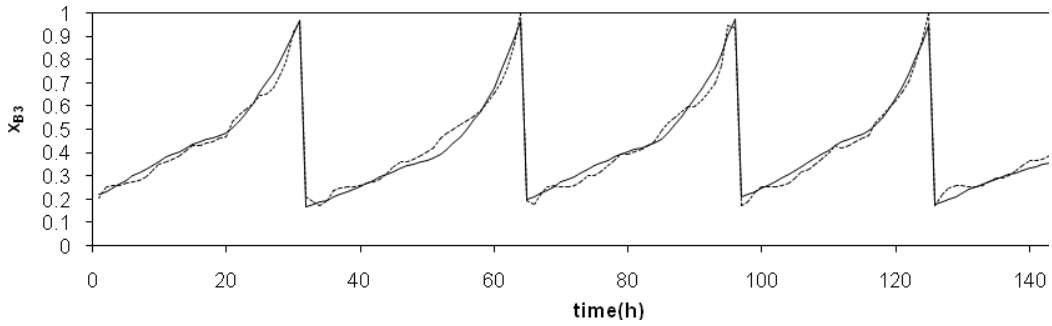


Fig. 2. Comparison between the actual values of the product composition (solid lines) and their estimates provided by OLS (dotted lines) for training set

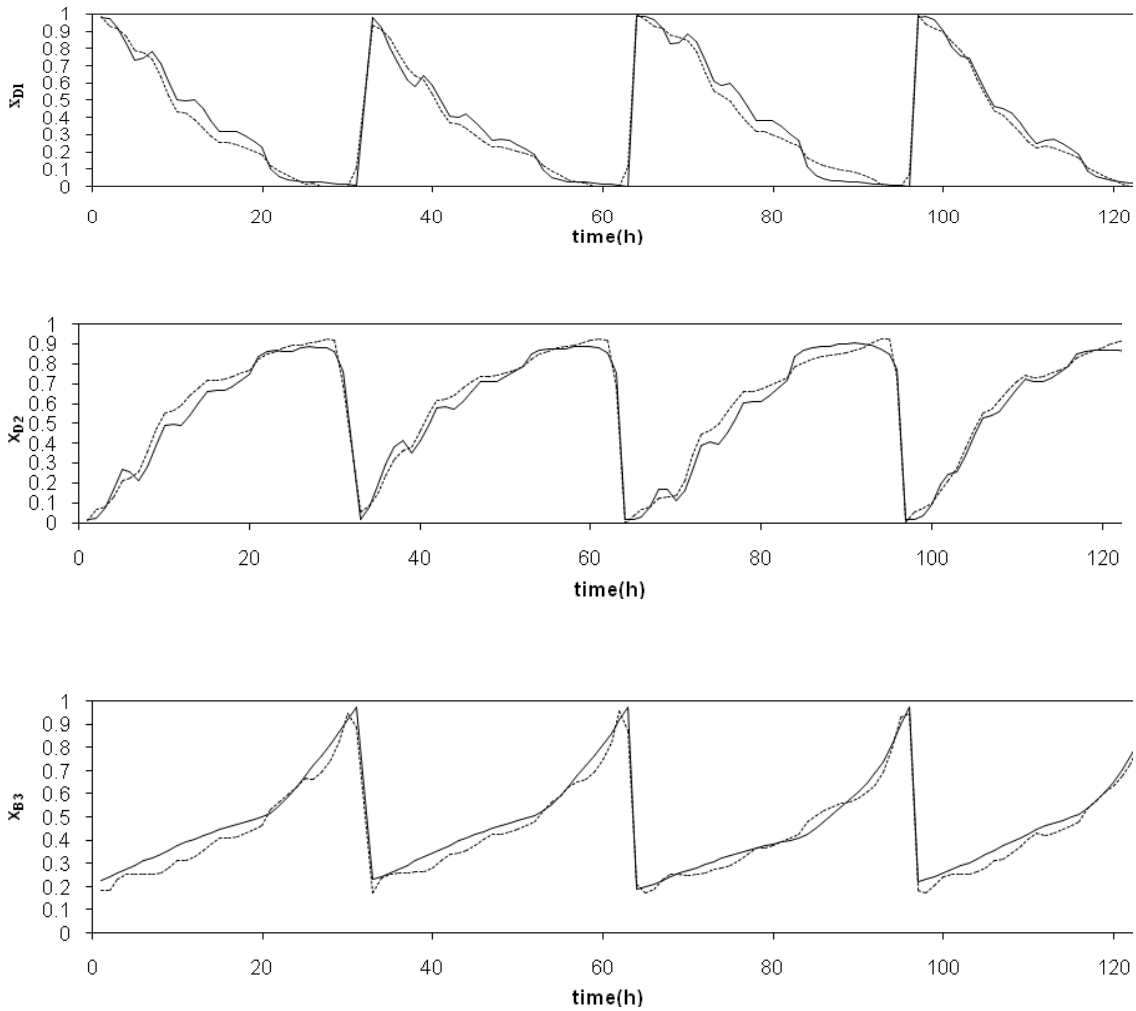


Fig 3. Comparison between the actual values of the product composition (solid lines) and their estimates provided by OLS (dotted lines) for test set

predict purity. The Gaussian function is used as basis functions to construct an RBF model with a uniform width σ = 0.90.

$$\phi(\mathbf{x}, c_i) = \exp\left(\frac{\|\mathbf{x}-c_i\|^2}{\sigma^2}\right) \quad (16)$$

where $\mathbf{x}$ is the network input vector, c_i are the center vectors. Fifteen points are used as candidate center set. The Orthogonal Least Square (OLS) subset selection algorithm was applied. During the forward regression model construction process, the model with 7 centers is automatically derived as the final model.

The results of the derived RBF model with 7 centers are shown in Fig 2 and Fig 3. The model RMSE of training and test is 0.0404 and 0.0460 respectively. It is can be seen that the model constructed by the OLS based RBF algorithm captured the underlying dynamics of the system very well.

6. CONCLUSION

In this paper, OLS based-RBF has been employed for estimating the distillate and reboiler composition from tray temperatures of batch distillation column. The results for multicomponent mixtures indicate that the estimator describe the process fairly accurately for varying reflux ratio.

The Orthogonal Least Square (OLS) algorithm is an efficient implementation of forward regression procedure for subset model selection. The Orthogonal Least Square (OLS) algorithm automatically selects the appropriate RBF centers from the training data and estimates the weights. Hence by using the proposed algorithm the complexity of the RBF network structure can be reduced. This will lead to faster and more efficient training.

As the estimation of composition rely on temperature measurements, any error in temperature measurement may lead to a wrong estimation of composition.

7. REFERENCES

1. Barolo, M, Pistillo, A., & Trotta, A. (2000). Issues in the development of a composition estimator for a middle vessel batch column. In E. F. Camacho, L. Basanez, J. A. de la Puente (Eds.),Advanced control of chemical processes 2000–IFAC ADCHEM 2000 (pp. 923–928). Oxford, UK: Elsevier.
2. Billings S.A., Chen S., and Korenberg M. J. (1989). Identification of MIMO non-linear systems using a forward-regression orthogonal estimator, International Journal of *Control*, 49, 2157–2189.
3. Chen, S., Billings S.A., and Luo W. (1989). Orthogonal least squares methods and their application to non-linear system identification, International Journal of Control, *50*(5),1873–1896.
4 .Chen, S., Billings, S.A., Cowan, C.F.N., & Grant, P.M.(1990). Non-linear systems identification using radial basis functions, International Journal of Systems Science, *21*, 2513-2539.
5. Chen, S., Cowan, C.F.N., and Grant, P.M.(1991). Orthogonal least squares learning algorithm for radial basis function networks, IEEE Transaction on Neural Networks, 2, 302-309.
6. Chen, S. (2002). Multi-output Regression Using a Locally Regularized Orthogonal Least-square Algorithm, IEE Proc. Vis. Image Signal Process, 149(4).
7. D.L. YU., J.B.Gomm., and D Williams.(1997) A Recursive Orthogonal Least Squares Algorithm for Training RBF Networks, Neural Processing Letters, 5,167–176.
8. E. Zamprogna E., Barlo M., Seborg D.E. (2004). Development of a soft sensor for a batch distillation column using linear and nonlinear PLS regression techniques, Control Engineering Practice, *12*, 917–929.
9. Jimenez L., Basualdo M. S., Gomez J. C., Toselli L., Rosa M. (2002). Nonlinear dynamic modeling of multicomponent batch

distillation: A case study, Brazilian Journal of Chemical Engineering, 19(3), 307-317.

10. Luyben, W.L. (1991). Process Modelling, Simulation and Control for Chemical Engineers, (2nd ed.) McGraw Hill, New York.
11. Mujtaba, I. M., Macchietto S. (1996) Simultaneous optimization of design and operation of multicomponent batch distillation column-single and multiple separation duties, Journal of Process Control, 6(1), 27-36.
12. Oisiovici, R. M., & Cruz, S. L. (2000). State estimation of batch distillation columns using an extended Kalman filter. Chemical Engineering Science, *55*, 4667–4680.
13. Osiovici, R. M., & Cruz, S. L. (2001). Inferential control of high-purity multicomponent batch distillation columns using an extended Kalman filter. Industrial and Engineering Chemistry Research, *40*, 2628–2639.
14. Quintero-Marmol, E., & Luyben, W. L. (1992). Inferential model based control of multicomponent batch distillation. Chemical Engineering Science, *47*, 887–898.
15. Quintero-Marmol, E., Luyben, W. L., & Georgakis, C. (1991).Application of an extended Luenberger observer to the control of multicomponent batch distillation. Industrial and Engineering Chemistry Research, *30*, 1870–1880.
16. Venkateswarlu, C., & Avantika, S. (2001). Optimal state estimation of multicomponent batch distillation. Chemical Engineering Science, *56*, 5771–5786.