Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

A semi-batch reactor modeling based on PWARX systems

Zeineb LASSOUED * **Kamel ABDERRAHIM** **

* *Numerical Control of Industrial Processes, National School of Engineers of Gabes, University of Gabes, St Omar Ibn-Khattab, 6029 Gabes, Tunisia.*
(e − mail : zeineb.lassoued1@gmail.com).
** *(e − mail : kamelabderrahim@yahoo.fr)*

Abstract: In this paper, we address the problem of identifying a semi-batch olive oil esterification reactor. In fact, this reactor can be considered as a PieceWise AutoRegressive eXogenous (PWARX) system. The Chiu's clustering procedure for the identification of PWARX systems is then applied. It consists in estimating both the parameter vector of each submodel and the coefficients of each partition. The results of the experimental validation illustrate the effectiveness of the proposed method. A comparative study with three existing approaches is also considered in this paper which shows that the proposed approach gives the best results in terms of precision.

Keywords: Hybrid systems, PWARX models, Identification, Chiu's clustering technique, Comparison study, Experimental validation.

1. INTRODUCTION

Batch reactors represent the cornerstone of several industrial plants such as chemical and pharmaceutical industries. In this paper, we consider a semi-batch reactor producing ester used in fine chemical industry like cosmetic products. This reactor prepares the ester by the reaction of a vegetable olive oil with free fatty acid and an alcohol at an equilibrium point with the elimination of water. Thus, the production of ester by this reactor includes three main steps which are heating, reacting and cooling. The first step allows to reach the equilibrium point which is obtained by increasing the temperature of the reaction products until a desired value. The purpose of the second step is to eliminate the resulting water of the reaction. This reactor uses a boiling technique to ensure this goal. Finally, the cooling step allows to decrease the temperature of the ester until a specified temperature. The optimization of the ester quality and the performance of the reactor can be guaranteed by an effective control system. The control of such reactor often requires that the dynamic behavior of the system must be represented by a mathematical model. This model can be built by the physical laws describing the heating and cooling phenomena that govern the behavior of the reactor. Consequently, it leads to a complicated nonlinear model. A solution to this problem consists in using the identification approach which allows to build a mathematical model from input-output data. The identification of the reactor is known to be a challenging identification problem due to the high nonlinearity, complexity, variability and uncertainty that such processes involve. Several approaches have been proposed in the literature for the identification of this reactor. These approaches can be classified in different ways depending on: the model structure, the solution type, the data processing strategy, etc.

The first approach is based on the linearization of the process around an operating point Mihoub et al. (2009a) Ben Abdennour et al. (2001). This approach allows the use of the linear systems identification and control methods which are highly developed in the literature. But, the obtained model is far from optimal. It is even quite surprising since it represents a nonlinear system by a single linear model describing the local behavior of the system around a single operating point.

To overcome this problem, the authors of Msahli et al. (2001) have suggested the use of the Volterra model. It is well known that the truncated Volterra model can represent any nonlinear system time invariant with fading memory Schetzen (1980) Mathews and Sicuranza (2000) Ogunfunmi (2007). Moreover, the parameters of this model are linearly related to the output, which allows the extension of some results of linear systems to nonlinear ones. However, it is important to note that the number of parameters of the Volterra model increases rapidly with the order and the memories. This is the main drawback of this model. Therefore, most of the researchers treat only the case where the order and memory are relatively small. This choice reduces the accuracy of the resulting model. The excitation signal which must be Gaussian represents also a fundamental problem for the identification of Volterra systems because in practical situations random signals may cause the wear of the actuators. Moreover, this method requires a very high number of measurements.

Another solution based on recurrent neural networks approach is proposed in Atig et al. (2010) Atig et al. (2012) to represent this reactor. The neural networks are capable of approximating any continuous nonlinear functions and have been applied to nonlinear process emulation. In fact, this approach consists in using a real time recurrent learning based on the gradient backpropagation learning algorithm. This algorithm is used for the adaptation of neural emulator parameters and weights. The major advantage of such method is that it doesn't depend on any preliminary knowledge about dynamics. But, it is well known that the gradient descent method is characterized by its slowness of convergence and it may converge to local minima.

Finally, we cite the approach proposed in Mihoub et al. (2009b)Ltaief et al. (2004). It consists in using the multimodel approach to represent this reactor. This approach is based on

the "divide to reign" strategy. It consists in decomposing the complex dynamic behavior of the system into a finite number of operating points. The set of local models is commonly known as the library of models. Each local model contributes partially to the global behavior of the system. Indeed, the output of the latter is obtained by aggregation of the contribution of each local model which is defined by a weighting function (also called validity). However, the construction of the library as well as the computation of validities is considered among the problems that are encountered in the multimodel approach.

In this paper, we consider the special case of multimodel where the validities of sub-models is binary. In fact, we preconize the use of the PieceWise Auto-Regressive eXogenous (PWARX) model to represent the reactor. It consists in using the notion of hybrid system which can be used to represent complex nonlinear systems. In fact, we can decompose the domain range of the non linear system into a set of operating regions Lin and Unbehauen (1992). For each one, a linear or affine model is associated. So, the considered complex non linear system becomes by modeling an hybrid system switching between linear or affine submodels Roll et al. (2004) Nakada et al. (2005) Wen et al. (2007) Xu et al. (2012). The alternative of considering the semi-batch reactor as a PWARX system seems to be very interesting because the characteristic of the system can be considered as linear in each operating step (heating, reacting, cooling).
This paper is organized as follows. Section II recalls the description of the process. In section III, the proposed method for the identification of the reactor is detailed. In section IV, the performance of the proposed approach is evaluated and compared with three existing methods through the experimental results. Section V concludes the paper.

2. PROCESS DESCRIPTION

The semi-batch reactor studied in this paper is an olive oil esterification reactor producing ester with a very high added value which is used in fine chemical industry such as cosmetic products. The esterification reaction between vegetable olive oil with free fatty acid and alcohol, producing ester, is given by the following equation:

$$Acide + Alcohol \leftrightarrow Ester + Water. \quad (1)$$

The ratio of the alcohol to acid represents the main factor of this reaction because the esterification reaction is an equilibrium reaction i.e. the reaction products, water and ester, are formed when equilibrium is reached. In addition, the yield of ester may be increased if water is eliminated from the reaction. The removal of water is achieved by the vaporisation technique while avoiding the boiling of the alcohol. In fact, we have used an alcohol (1- butanol), characterized by a boiling temperature of 118° C which is greater than the boiling temperature of the water (close to 100° C) . In addition, the boiling temperatures of the fatty acid (oleic acid) and the ester are close to 300°C. Therefore, the boiling point of water may be provided by a temperature slightly greater than 100° C.
The reactor is constituted essentially of:

- A reactor with double-jackets: it has a cylindrical shape manufactured in stainless steel. It is equipped with a bottom valve for emptying the product, an agitator, an orifice introducing the reactants, a sensor of the reaction mixture temperature, a pressure sensor and an orifice for the condenser. The double-jackets ensure the circulation of a coolant fluid which is intended for heating or for cooling the reactor.
- A heat exchanger: It allows to heat or to cool the coolant fluid circulating through the reactor jacket. Heating is carried out by three electrical resistances controlled by a dimmer for varying the heating power. It is intended to achieve the required reaction temperature of the esterification. Cooling is provided by circulating cold water through the heat exchanger. It is used to cool the reactor when the reaction is completed.
- A condenser: It allows to condense the steam generated during the reaction. It plays an important role because it is also used to indicate the end of the reaction which can be deduced when no more water is dripping out of the condenser.
- A data acquisition card between the reactor and the calculator.

The block diagram of the process is shown in Fig.1.

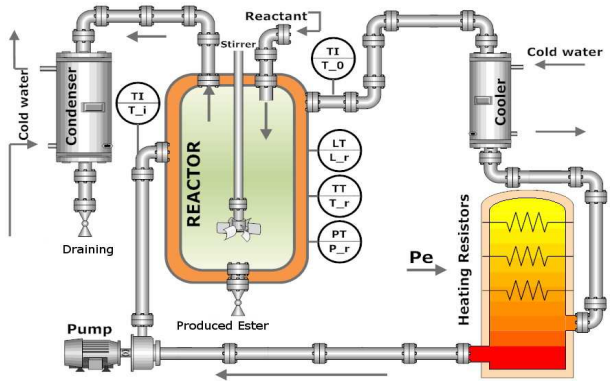

Fig. 1. Block diagram of the reactor.

The ester production by this reactor is based on three main steps as illustrated in Fig.2

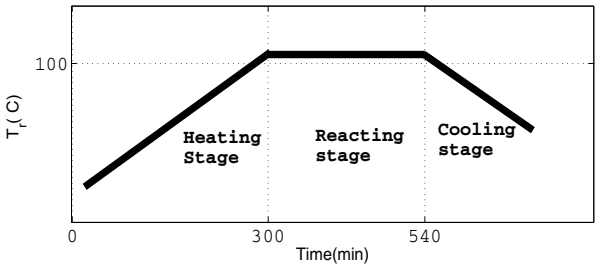

Fig. 2. Specific trajectory of the reactor temperature.

- Heating stage: The reactor's temperature T_r is increased to 105° C.
- Reacting stage: The temperature should be kept constant at 105° C until the end of the reaction which can be detected by the absence of water at the condenser (when no more water is dripped out of the condenser).
- Cooling stage: The reactor's temperature is decreased.

An experimental study carried out on the reactor showed that the reactor can be considered as a Single-Input Single-Output (SISO) plant. The heating power P_e and the reaction temperature T_r represent respectively the input and the output of the

system. However, the variation of the quality of the reactants inside the reactor as well as the external effects can be considered as random disturbances. Moreover, from the same study, it has been noted that the reactor is non linear. In fact, the static characteristic of the system as shown in Fig.3 confirms the non linearity.

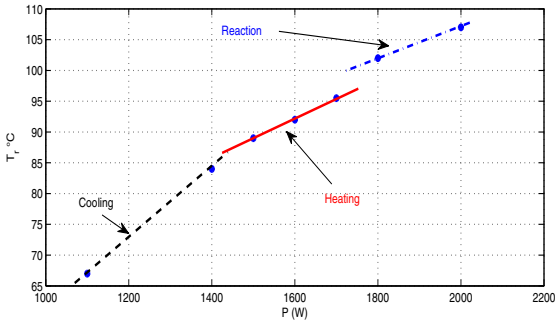

Fig. 3. Static characteristic of the reactor.

3. THE PROPOSED APPROACH FOR MODELING THE REACTOR

Based on the results illustrated in Fig.3, we remark that each step of the operating cycle presents a linear behavior. Consequently, this reactor can be modeled by a PWARX model.

3.1 Model and assumptions

We propose the following PWARX model to represent this reactor:

$$y(k) = f(\varphi(k)) + e(k). \tag{2}$$

where
- $y(k) \in R$ is the system output (k is the now time index),
- $e(k)$ is assumed to be a Gaussian sequence independent and identically distributed with zero mean and finite variance σ^2,
- $\varphi(k)$ is the vector of regressors which belongs to a bounded polyhedron H in R^d:

$$\varphi(k) = [y(k-1),\ldots,y(k-n_a)\, u(k-1),\ \ldots,u(k-n_b)]^T. \tag{3}$$

where $u(k) \in R$ is the system inputs, n_a and n_b are the system orders assumed to be known and $d = n_a + n_b + 1$.
- f is a piecewise affine function defined by:

$$f(\varphi) = \begin{cases} \theta_1^T \bar{\varphi} & if \quad \varphi \in H_1 \\ \vdots & \\ \theta_s^T \bar{\varphi} & if \quad \varphi \in H_s \end{cases} \tag{4}$$

where $\bar{\varphi} = \begin{bmatrix} \varphi^T & 1 \end{bmatrix}^T$, s is the number of sub-models supposed to be unknown, $\theta_i \in R^{d+1}$ are the parameter vectors. $\{H_i\}_{i=1}^s$ are polyhedral partitions of the bounded domain H verifying:

$$\begin{cases} \bigcup_{i=1}^{s} H_i = H \\ H_i \bigcap H_j = \emptyset, \quad \forall i \neq j \end{cases} \tag{5}$$

3.2 The used identification method

The decomposition of the state-input domain into a finite number of non-overlapping convex polyhedral regions and the association of a simple linear or affine model to each region represent the main steps of the construction of PWARX models. Consequently, the PWARX identification problem involves both the estimation of the parameters of the submodels and the hyperplanes defining the partitions of the state-input regression. Numerous methods have been proposed in the literature for the identification of PWARX models. These methods can be classified into several categories of solutions such as the algebraic solution Tian et al. (2011), the clustering-based solution Ferrari-Trecate et al. (2003), the Bayesian solution Juloski et al. (2005), the bounded-error solution Bemporad et al. (2005), the sparse optimization solution Laurent (2011) etc. Only the clustering based procedure is considered in this paper. Most of these methods are based on classical clustering algorithms such as k-means methods Ferrari-Trecate et al. (2003) Ferrari-Trecate et al. (2001). The performance of these algorithms is often degraded in the case of poor initialization because they consist in minimizing a non linear criterion for which local minima may exist. Moreover, the outliers are not removed effectively. Also, most of these algorithms assume that the number of submodels is a priori known. To overcome these problems, we have proposed an alternative solution to these methods based on the Chiu's clustering algorithm Lassoued and Abderrahim (2013b) Lassoued and Abderrahim (2013a). The main steps of this approach can be summarized as follows.

Constructing new data set from the initial data set For every pair of data $\{\varphi(k), y(k)\}_{k=1}^N$, we construct a local set C_k collecting $\{\varphi(k), y(k)\}$ and its $(n_\rho - 1)$ nearest neighbors satisfying:

$$\forall(\breve{\varphi},\breve{y}) \in C_k, \left\|\varphi(k) - \breve{\varphi}\right\|^2 \leq \|\varphi(k) - \hat{\varphi}\|^2, \forall(\hat{\varphi},\hat{y}) \notin C_k. \tag{6}$$

The parameter n_ρ is chosen randomly as $n_\rho > d+1$. This parameter decisively influences on the performance of the algorithm. We define ρ_k as follows: $\rho_k = \left\{t_k^1, \cdots, t_k^{n_\rho}\right\}$. It contains, in ascending order, the indexes of the elements belonging to C_k.

Estimating a parameter vector for each data set For each local set C_k, we identify an affine model using least square method to determine the local parameters θ_k:

$$\theta_k = (\phi_k^T \phi_k)^{-1} \phi_k^T Y_k. \tag{7}$$

where

$$\phi_k = \left[\bar{\varphi}(t_k^1)...\bar{\varphi}(t_k^{n_\rho})\right]^T, \quad Y_k = \left[y(t_k^1)...y(t_k^{n_\rho})\right]^T.$$

Classifying the parameter vectors The obtained parameter vectors $\{\theta_k\}_{k=1}^N$ are classified into s disjoint clusters using a suitable classification technique. The proposed method for data classification is based on the use of Chiu's clustering technique Lassoued and Abderrahim (2013b) Lassoued and Abderrahim (2013a) which can be summarized by the following algorithm:

- Dispose of $\{\theta_i\}_{i=1}^N$ from a given data set (φ_i, y_i).
- Compute P_i for every $\{\theta_i\}_{i=1}^N$according to the following equation.

$$P_i = \sum_{j=1}^{N} e^{-\frac{4}{r_a^2}\|\theta_i - \theta_j\|^2} \tag{8}$$

- Determine the filtered data points $\{\theta_i\}_{i=1}^{N'}$, $(N' < N)$
- Compute the first cluster center θ_1^* from (8)
- **repeat**
  Compute the other cluster centers according to the updated potential formula (9)

$$P_i \Leftarrow P_i - P_k^* e^{-\frac{4}{r_b^2}\|\theta_i - \theta_k^*\|^2} \tag{9}$$

if $P_k^* > \gamma$ **then**
Compute $V(c)$ such as:

$$V(c) = \|\theta_k^* - \theta_c^*\|, \quad c = 1,...,k-1 \tag{10}$$

where $\theta_k{}^*$ is the current cluster center and θ_c^*, $c = 1,...,k-1$ are the last selected ones.
if $V(c) > \varepsilon$, $c = 1,...,k-1$ **then**
accept $\theta_k{}^*$ as a cluster center and continue
else reject $\theta_k{}^*$ and compute a new potential
end if
else reject $\theta_k{}^*$ and break
end if
until $V(c) \leq \varepsilon, \quad c = 1,...k-1$

While ε is a small parameter characterizing the minimum distance between the new cluster center and the existing ones.

In algorithm 1, we compute the potential P_i of every local parameter vector $\{\theta_i\}_{i=1}^{N}$. The first cluster center has the highest value of potential. Then, the next cluster center take the highest remaining potential value in condition that the new center is far enough from the previously computed center. ε is then the small parameter characterizing the minimum distance between the new cluster center and the existing ones. r_a is the radius defining the neighborhood of data points. The parameter r_b is a positive constant that must be chosen larger than r_a to avoid obtaining closely spaced cluster centers.

Estimating the parameters of each sub-model After classifying the data, we can determine the s ARX sub-models defined by the parameter vectors $\theta_i, \quad i = 1,...,s$ by using the least square method.

Estimating the regions The final step consists in determining the regions H_i. The methods of statistical learning such as the Support Vector Machines (SVM) offer an interesting solution to accomplish this task Wang (2005) R.O.Duda et al. (2001). The SVM are a popular learning method for classification, regression and other learning tasks Hsu and Lin (2002) Weston and Watkins (1999).
In our case, it is a matter of finding for every $i \neq j$ the hyperplane that separates points existing in H_i and in H_j. Given two sets H_i and H_j, $i \neq j$, the linear separation problem is to find $w \in R^d$ and $b \in R$ such that:

$$\begin{aligned} w^T \varphi_k + b > 0 \ \ \forall \ \varphi_k \in H_i, \\ w^T \varphi_k + b < 0 \ \ \forall \ \varphi_k \in H_j. \end{aligned} \tag{11}$$

This problem can be easily rewritten as a feasibility problem with linear inequality constraints. The estimated hyperplane separating H_i from H_j is denoted with $M_{i,j}\varphi = m_{i,j}$ where $M_{i,j}$ and $m_{i,j}$ are matrices of suitable dimensions. Moreover, we assume that the points in H_i belong to the half-space $M_{i,j}\varphi \leq m_{i,j}$.
The regions $\{H_i\}_{i=1}^{s}$ are obtained by solving these linear inequalities Ferrari-Trecate et al. (2003):

$$\left[M'_{i,1}...M'_{i,s}M'\right] \bar{\varphi} \leq \left[m'_{i,1}...m'_{i,s}m'\right]. \tag{12}$$

where $Mx \leq m$ are the linear inequalities describing the bounded domain H.

4. IDENTIFICATION RESULTS

Some input-output measurements are picked out from the reactor in order to determine a model to this process. We have taken two measurement files, one for the identification and another one for the validation.
In fact, a pseudo-random, binary sequence is applied as input to the real system with three different amplitude ranges in order to excite the three steps of the operating cycle of the reactor (heating, reacting and cooling). This latter is applied to the reactor with a sampling time equal to $180s$. The obtained measurements presented in Fig. 4 are then used to identify the reactor with different methods.

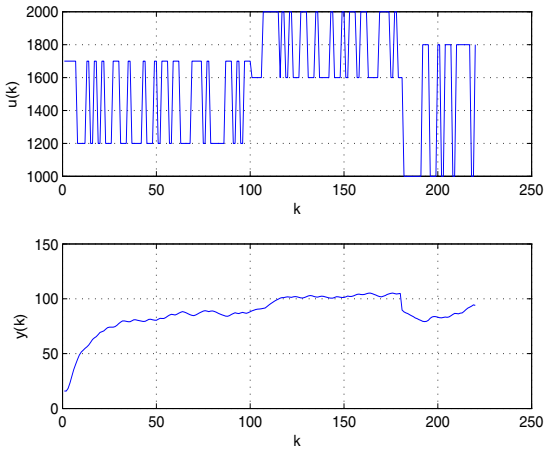

Fig. 4. The real input-output evolution.

4.1 Results of the proposed method

Previous works have demonstrated that the adequate estimated orders n_a and n_b of each submodel are equal to two. Thus, we can adopt the following structure:

$$y(k) = \begin{cases} -a_{1,1}y(k-1) - a_{1,2}y(k-2) + b_{1,1}u(k-1) + \\ b_{1,2}u(k-2) \quad if \quad \varphi(k) \in H_1 \\ \vdots \\ a_{s,1}y(k-1) + a_{s,2}y(k-2) + b_{s,1}u(k-1) + \\ b_{s,2}u(k-2) \quad if \quad \varphi(k) \in H_s \end{cases} \tag{13}$$

where the regressor vector is defined by:

$$\varphi(k) = [-y(k-1), -y(k-2), u(k-1), u(k-2)]^T$$

and the parameter vectors are denoted by:

$$\theta_i(k) = [a_{i,1}, a_{i,2}, b_{i,1}, b_{i,2}], \ i = 1,...,s.$$

We apply the proposed identification procedure in order to represent the reactor by a PWARX model with a number of neighboring $n_\rho = 70$. Our purpose is to estimate the number of sub-models s, the parameter vectors $\theta_i(k), \ i = 1,...,s$ and the hyperplanes defining the partitions $\{H_i\}_{i=1}^{s}$.
The obtained results are as follows:

- The number of sub-models is $s = 3$.
- The parameter vectors $\theta_i(k)$, $i = 1, 2$ *and* 3 are illustrated in Table 1.

Table 1. Estimated parameter vectors

Parameter vectors	Estimated parameters
θ_1	$\begin{bmatrix} -1.4256 \\ 0.4508 \\ 4.9853 \quad 10^{-4} \\ 0.0010 \end{bmatrix}$
θ_2	$\begin{bmatrix} -1.1604 \\ 0.2111 \\ 0.0015 \\ 0.0014 \end{bmatrix}$
θ_3	$\begin{bmatrix} -1.0847 \\ 0.1490 \\ -3.9782 \quad 10^{-4} \\ 0.0040 \end{bmatrix}$

- The partitions $\{H_i\}_{i=1}^{s}$ which are defined by the following inequalities:

$$H_i = \left\{ \varphi \in \Re^4 : M_i \bar{\varphi} \geq 0 \right\}. \quad (14)$$

$$M_1 = \begin{bmatrix} -0.3869 & 0.3130 & 0.0001 - 0.0015 & -4.8740 \\ -0.0164 & 0.0035 & 0.0018 - 0.0026 & 0.7355 \end{bmatrix}. \quad (15)$$

$$M_2 = \begin{bmatrix} 0.4548 - 0.3862 & -0.0001 & 0.0014 & 4.5891 \\ -0.2871 & 0.2519 & 0.0021 - 0.0034 & -0.4359 \end{bmatrix}. \quad (16)$$

$$M_3 = \begin{bmatrix} 0.0163 - 0.0034 - 0.0018 & 0.0026 - 0.7335 \\ 0.1342 - 0.1041 - 0.0019 & 0.0031 \quad 0.2969 \end{bmatrix}. \quad (17)$$

The obtained output is presented in Fig.5.

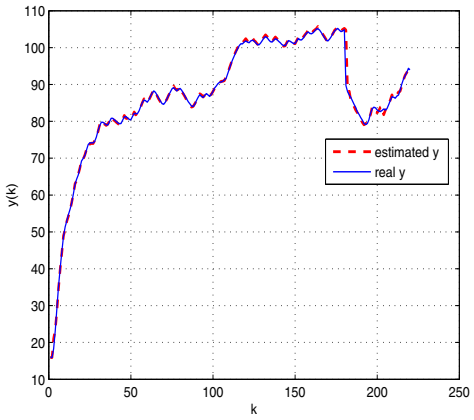

Fig. 5. The real and the estimated outputs of the proposed method.

To validate the obtained models, we have considered a new input-output measurement file. The obtained results presented in Fig.6 prove the efficiency of the proposed method.

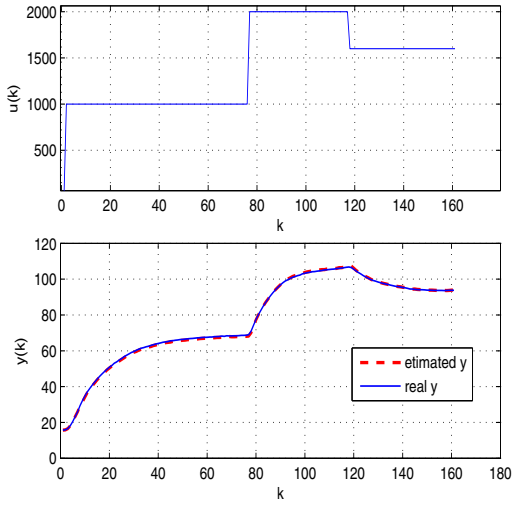

Fig. 6. The input, the real and the estimated validation outputs.

4.2 Comparison with the existing methods

In this section, our proposed method is compared with three existing methods: the classic method Ben Abdennour et al. (2001), the multimodel method Mihoub et al. (2009b)Ltaief et al. (2004)and the neural emulator method Atig et al. (2012). The system's output of every approach is presented in Fig.7.

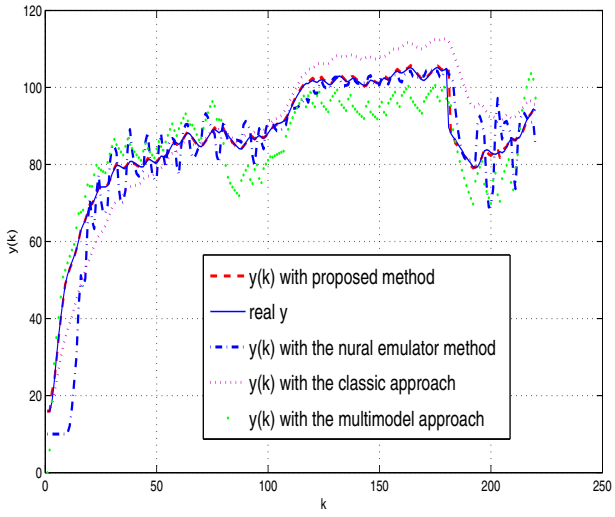

Fig. 7. The real and the estimated outputs.

Based on the results presented in Fig.7, we observe that:

- The linear approach gives poor results which can be justified by the fact that it represents a nonlinear plant by a linear model.
- The recurrent neural networks deliver estimated output with large error. These poor results are due to the slowness of convergence of the gradient back propagation learning algorithm.
- The multimodel poorly performs by comparison to the proposed approach. The reason is that the computation of validities is based on the assumption that the sum of all validity is equal to 1. This assumption raises problems

in practice situations when only one model is valid. In fact, all methods of calculating validities proposed in the literature affect a validity close to 1 for the valid submodel and validity close to 0 for the other submodels. These small validities introduce errors on the global system output.

- The proposed method gives the best estimation results which prove the importance of representing the reactor by a PWARX model.

5. CONCLUSIONS

In this paper, we have considered the problem of identification of a semi-batch reactor producing ester by the reaction of a vegetable olive oil with free fatty acid and an alcohol. In fact, we have recalled the main steps used by this reactor to produce ester. The static characteristic of the reactor is also presented using input-output measurements. This presentation shows that the behavior of each step (heating, boiling and cooling) can be approximated by a linear model. Consequently, we have suggested the use of PWARX models to represent this reactor. To implement this solution, we have proposed the use of the Chiu's clustering algorithm. The obtained results have shown the performance of the proposed solution by comparison with the main existing approaches. In the next work, we propose the control of this reactor using the obtained PWARX model.

REFERENCES

Atig, A., Druaux, F., D.Lefebvre, K.Abderrahim, and Abdennour, R.B. (2012). Adaptive control design using stability analysis and tracking errors dynamics for nonlinear square mimo systems. *Engineering Applications of Artificial Intelligence, Elsevier*, 25, 1450–1459.

Atig, A., F.Druaux, D.Lefebvre, K.Abderrahim, and Ben.Abdennour, R. (2010). Neural emulator and controller with decoupled adaptive rates for nonlinear systems: application to chemical reactors. *International Journal on Sciences and Techniques of Automatic control and Computer Engineering, IJ-STA*, 4(2), 1298–1319.

Bemporad, A., Garulli, A., Paoletti, S., and Vicino, A. (2005). A bounded-error approach to piecewise affine system identification. *IEEE Transactions on Automatic Control*, 50(10), 1567–1580.

Ben Abdennour, R., Borne, P., Ksouri, M., and M'sahli, F. (2001). Identification et commande numrique des procédés industriels. *Editions Techip*.

Ferrari-Trecate, G., Muselli, M., Liberati, D., and Morari, M. (2001). Identification of piecewise affine and hybrid systems. In *Proceedings of American Control Conference, Vol. 5, pp. 3521-3526*.

Ferrari-Trecate, G., Muselli, M., Liberati, D., and Morari, M. (2003). A clustering technique for the identification of piecewise affine systems. *Automatica*, 39(2), 205–217.

Hsu, C. and Lin, C. (2002). A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, 13(2), 415–425.

Juloski, A., Weiland, S., and Heemels, W. (2005). A bayesian approach to identification of hybrid systems. *IEEE Transactions on Automatic Control*, 50(10), 1520–1533.

Lassoued, Z. and Abderrahim, K. (2013a). New approaches to identification of pwarx systems. *Mathematical Problems in Engineering, Hindawi Publishing Corporation*.

Lassoued, Z. and Abderrahim, K. (2013b). A new clustering technique for the identification of pwarx hybrid models. *The 9^{th} Asian Control Conference ASCC, Istanbul, Turkey*.

Laurent, B. (2011). Identification of switched linear systems via sparse optimization. *Automatica*, 47(4), 668–677.

Lin, J. and Unbehauen, R. (1992). Canonical piecewise-linear approximations. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 39(8), 697–699.

Ltaief, M., Abderrahim, K., Ben Abdennour, R., and Ksouri, M. (2004). Systematic determination of a models' base for the multimodel approach: experimental validation. *WSEAS Transactions on Electronics*, 1(2), 331–336.

Mathews, V.J. and Sicuranza, G.L. (2000). Polynomial signal processing. *Wiley*.

Mihoub, M., Nouri, A.S., and Ben.Abdennour, R. (2009a). Real-time application of discrete second order sliding mode control to a chemical reactor. *Control Engineering Practice*, 17(9), 1089–1095.

Mihoub, M., Nouri, A.S., and Ben.Abdennour, R. (2009b). A real time application of discrete second order sliding mode control to a semi-batch reactor: a multimodel approach. *International Journal of Modelling, Identification and Control*, 6(2), 156–163.

Msahli, F., Ben.Abdennour, R., and M.Ksouri (2001). Identification and predictive control of a nonlinear process using a parametric volterra model. *International Journal of Computational Engineering Science*, 2(04), 633–651.

Nakada, H., Takaba, K., and Katayama, T. (2005). Identification of piecewise affine systems based on statistical clustering technique. *Automatica*, 41(5), 905–913.

Ogunfunmi, T. (2007). Adaptive nonlinear system identification: The volterra and wiener model approaches. *Springer*.

R.O.Duda, P.E.Hart, and D.G.Stork (2001). Pattern classification. *Second edition, John Wiley and Sons*.

Roll, J., Bemporad, A., and Ljung, L. (2004). Identification of piecewise affine systems via mixed-integer programming. *Automatica*, 40(1), 37–50.

Schetzen, M. (1980). The volterra and wiener theories of nonlinear systems. *John Wiley and Sons, Inc., New-York*.

Tian, Y., Floquet, T., Belkoura, L., Perruquetti, W., et al. (2011). Algebraic switching time identification for a class of linear hybrid systems. *Nonlinear Analysis, Hybrid Systems*, 5(2), 233–241.

Wang, L. (2005). Support vector machines : Theory and applications. *Springer*.

Wen, C., Wang, S., Jin, X., and Ma, X. (2007). Identification of dynamic systems using piecewise-affine basis function models. *Automatica*, 43(10), 1824–1831.

Weston, J. and Watkins, C. (1999). Support vector machines for multi-class pattern recognition. In *Proceedings of the Seventh European Symposium on Artificial Neural Networks*, volume 4, 219–224.

Xu, J., Huang, X., Mu, X., and Wang, S. (2012). Model predictive control based on adaptive hinging hyperplanes model. *Journal of Process Control, Elsevier*, 22(10), 1821–1831.