

PWARX Model Identification Based on Clustering Approach

Zeineb Lassoued and Kamel Abderrahim

Abstract This chapter addresses the problem of clustering based procedure for the identification of PieceWise Auto-Regressive eXogenous (PWARX) models. In order to overcome the main drawbacks of the existing methods such as their sensitivity to poor initializations and the existence of outliers, we propose the use of the Chiu's clustering algorithm and the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. A comparative study of the two proposed approaches with the k-means method is achieved in simulation. The results of experimental validation are also presented to illustrate the effectiveness of the proposed methods.

Keywords Identification · PWARX systems · Clustering approach · Chiu's algorithm · DBSCAN algorithm · Experimental validation

1 Introduction

Hybrid systems are heterogeneous dynamical systems that arise out of the interaction of continuous and discrete dynamics. The continuous behavior is the fact of the natural evolution of the physical process whereas the discrete behavior can be due to the presence of switches, operating phases, transitions, computer program codes, etc. These hybrid dynamics characterize the behavior of a broad class of physical systems, for example, the real-time control systems where physical processes are controlled by embedded controllers. The notion of hybrid system can also be used to represent complex nonlinear continuous systems. In fact, the operating range of a nonlinear system can be decomposed into a group of operating

Z. Lassoued · K. Abderrahim (✉)

Numerical Control of Industrial Processes, National School of Engineers of Gabes,
University of Gabes, St. Omar Ibn-Khattab, 6029 Gabes, Tunisia
e-mail: kamelabderrahim@yahoo.fr

Z. Lassoued

e-mail: zeineb.lassoued1@gmail.com

© Springer International Publishing Switzerland 2015

Q. Zhu and A.T. Azar (eds.), *Complex System Modelling and Control Through Intelligent Soft Computations*, Studies in Fuzziness and Soft Computing 319,
DOI 10.1007/978-3-319-12883-2_6

165

point. For each operation point, we associate a simple sub-model (linear or affine) with it. Indeed, a complex system can be modeled as a hybrid system switching between simple sub-models.

This chapter addresses the problem of identification of hybrid systems represented by piecewise autoregressive models with exogenous input (PWARX). This problem consists in building mathematical models of hybrid systems from observed input-output data. The PWARX models have attracted a considerable attention in recent years, since they provide an efficient solution for modeling a wide range of engineering applications (Roll et al. 2004; Nakada et al. 2005; Wen et al. 2007; Xu et al. 2012). In addition, these models are able to approximate any nonlinear system with arbitrary accuracy (Lin and Unbehauen 1992). Moreover, the PWA model can be considered as a generic representation for other hybrid models such as jump linear models (JL models) (Vidal et al. 2002), Markov jump linear models (MJL models) (Doucet et al. 2001), mixed logic dynamical models (MLD models) (Bemporad et al. 2000), max-min-plus-scaling systems (MMPS models) (De Schutter and Van den Boom 2000), linear complementarity models (LC models) (Vander-Schaft and Schumacher 1998), extended linear complementarity models (ELC models) (De Schutter and De Moor 1999). In fact, the transfer of the results of PWARX models to other classes of hybrid systems is insured thanks to the properties of equivalence of PWARX models (Heemels et al. 2001). The PWARX models are obtained by decomposing the regression domain into a finite number of non-overlapping convex polyhedral regions and by associating a simple linear model with each region. Consequently, two main problems must be considered for the identification of PWARX models: one is the estimation of the parameters of the sub-models and two is the determination of the hyperplanes defining the partitions of the state-input regression. Consequently, the identification of PWARX models is one of the most difficult problems that represent an area of research where considerable work has been done in the last decade. In fact, numerous solutions have been proposed in the literature for the identification of the PWARX models such as 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 greedy solution (Bemporad et al. 2003), the sparse optimization solution (Bako 2011; Bako and Lecoeuche 2013), and so on. The sparse solutions do not smooth out the effect of the measurement noise. Then, they often fail in real time applications since the measurement data are usually contaminated by an unknown additional noise. The greedy algorithms are very time consuming since they involve the solution of NP-hard problems. In addition, it can cause a loss of information because it sometimes fails to associate data to the appropriate regressors. The Bayesian approach assumes that the probability density functions of the unknown parameters of the system are known a priori. Otherwise, it requires an additional sequential processing to improve the identification results. The clustering solution is based on a simple and instructive procedure. It does not require a priori knowledge of the system. Therefore, only the clustering approach is considered in this chapter. This solution consists of three main steps, which are data classification, parameter estimation and region reconstruction. It is easy to remark that the

performance of this approach depends on the efficiency of the used classification algorithm (Lassoued and Abderrahim 2013a, b, c, d, 2014a, b). The early methods have favored the simplicity of implementation. In fact, they present several drawbacks, which can be summarized as follows:

- Most of them are based on the optimization of nonlinear criteria. Consequently, they may converge to local minima in the case of poor initializations.
- Their performances degrade in the case of the presence of outliers in the data to be classified.
- Most of them assume that the number of sub-models is a priori known.

To overcome these problems, we have proposed the use of other clustering algorithms such as Chiu’s method (Chiu 1997) and Density Based Spatial Clustering of Applications with Noise (DBSCAN) method (Chaitali 2012; Sander et al. 1998). This choice is justified by the fact that these algorithms automatically generate the number of models. In addition, they are characterized by their robustness to the classification of noisy measurements that containing also outliers.

This chapter is organized as follows. Section 2 presents the assumptions for PWARX model identification. In Sect. 3, we recall the main steps of the identification of PWARX systems based on clustering algorithm and its main drawbacks. Section 4 proposes two solutions to overcome the main problems of the existing methods. In Sect. 5, we present three simulation examples in order to illustrate the performance of the proposed solutions and to compare their efficiency with the modified k-means method. Section 6 proposes an application of the developed approach to an olive oil esterification reactor.

2 Piecewise Affine System Identification

Consider a discrete-time PieceWise Auto-Regressive eXogenous model (PWARX) with input $u(k) \in \mathbb{R}$, output $y(k) \in \mathbb{R}$ defined in the bounded polyhedron regressor space $H \subset \mathbb{R}^d$ ($d = n_a + n_b + 1$). The system is decomposed in s different modes $\{H_i\}_{i=1}^s$, in each one an ARX model is associated:

$$y(k) = f(\varphi(k)) + e(k). \tag{1}$$

f is a piecewise affine function defined by:

$$f(\varphi) = \begin{cases} \theta_1^T \bar{\varphi} & \text{if } \varphi \in H_1 \\ \vdots & \\ \theta_s^T \bar{\varphi} & \text{if } \varphi \in H_s \end{cases} \tag{2}$$

where

$$\bar{\varphi} = [\varphi^T \quad 1]^T. \quad (3)$$

$e(k)$ is the additive noise and $\varphi(k)$ is the regressor vector, containing past input and output observations, defined as:

$$\varphi(k) = [y(k-1) \dots y(k-n_a) \quad u(k-1) \dots u(k-n_b)]^T. \quad (4)$$

$\theta_i \in \mathbb{R}^{d+1}$ is the parameter vector, valid in H_i , defined as follows:

$$\theta_i^T = [a_1 \quad a_2 \quad \dots \quad a_{n_a} \quad b_1 \quad b_2 \quad \dots \quad b_{n_b} \quad g] \quad (5)$$

where a_i and b_i are the coefficients of the model related respectively to the output and the input data, while n_a and n_b are the model orders. g is the independent affine coefficient.

Problem statement

Given input-output data generated by a PWARX system, we are interested simultaneously in identifying the number of submodels s , the parameter vectors $\{\theta_i\}_{i=1}^s$ and the partitions $\{H_i\}_{i=1}^s$ taking into account the following assumptions:

- The orders n_a and n_b of the system are known.
- The noise $e(k)$ is assumed to be a Gaussian process independent and identically distributed with zero mean and finite variance σ^2 .
- The regions $\{H_i\}_{i=1}^s$ are the polyhedral partitions of a bounded domain $H \subset \mathbb{R}^d$ such that:

$$\begin{cases} \bigcup_{i=1}^s H_i = H \\ H_i \cap H_j = \emptyset \quad \forall i \neq j \end{cases} \quad (6)$$

3 Clustering Based PWARX Identification

The main steps of the clustering-based approach for the identification of PWARX models can be summarized as follows: constructing small data set from the initial data set, estimating a parameter vector for each small data set, classifying the parameter vectors in s clusters, classifying the initial data set and estimating the s sub-models with their partitions.

1. Form $\{\varphi(k), y(k)\}_{k=1}^N$ from the given dataset $S = (u(k), y(k))$, $k = 1, \dots, N$
2. Create local datasets C_k and identify the local parameter vectors θ_k

- (a) Choose n_ρ , the cardinality of data points to be contained in C_k , randomly.
- (b) For each dataset $\varphi(k), y(k)$, build C_k containing $\{\varphi(k), y(k)\}$ and its $(n_\rho - 1)$ nearest neighbors satisfying:

$$\left\| \varphi(k) - \tilde{\varphi} \right\|^2 \leq \left\| \varphi(k) - \hat{\varphi} \right\|^2, \quad \forall (\hat{\varphi}, \hat{y}) \notin C_k. \quad (7)$$

- (c) Determine θ_k for each data in $C_k, k = 1, \dots, N$ using the least square method.

$$\theta_k = (\phi_k^T \phi_k)^{-1} \phi_k^T Y_k. \quad (8)$$

where

$$\phi_k = [\bar{\varphi}(t_k^1) \dots \bar{\varphi}(t_k^{n_\rho})]^T,$$

$$Y_k = [y(t_k^1) \dots y(t_k^{n_\rho})]^T.$$

and $(t_k^1, \dots, t_k^{n_\rho})$ are the indexes of the elements belonging in C_k

3. Cluster the local parameter vectors $(\theta_k, k = 1, \dots, N)$ into s disjoint clusters while determining the value of s by using a suitable classification technique.
4. Identify the final models $\{\theta_i\}_{i=1}^s$.
5. Estimate the polyhedral partitions $\{H_i\}_{i=1}^s$ i.e. estimate the hyperplanes separating H_i from $H_j, i \neq j$. This is a standard pattern recognition/classification problem that can be solved by several established techniques. The most common technique is the Support Vector Machines (SVM) (Wang 2005; Duda et al. 2001).

The classification of data represents the main step for PWARX system identification because a successful identification of models' parameters and hyperplanes depends on the correct data classification. For the sake of simplicity, the early approaches use classical clustering algorithms for the data classification such as k-means algorithms.

However, these algorithms present several drawbacks. In fact, they may converge to local minima in the case of poor initializations because they are based on the minimization of non linear criterion. Furthermore, their performances degrade in the case of the presence of outliers in the data to be classified. In addition, most of them assume that the number of sub-models is a priori known.

4 The Proposed Clustering Techniques

In order to improve the identification results we propose the use of other classification algorithms such as Chiu's algorithm and DBSCAN algorithm.

4.1 The Chiu's Clustering Technique

The Chiu's clustering method is a modified form of the Mountain method for cluster estimation (Chiu 1994). Each data point is considered as a potential cluster center instead of considering it as a grid point. This method is very advantageous compared with the Mountain method:

- The number of points to be evaluated is equal to the number of data points.
- It does not need to specify a grid solution which trades off between the accuracy and the computational complexity.
- It improves the computational efficiency and robustness of the original method.

Chiu's classification method consists in computing a potential value for each point of the data set based on its distances to the other data points and consider each data point as a potential cluster center. The point having the highest potential value is chosen as the first cluster center. The key idea in this method is that once the first cluster center is chosen, the potential of all other points is reduced according to their distance from the cluster center. All the points which are close to the first cluster center will have greatly reduced potentials. The next cluster center take then the highest remaining potential value. The procedure for determining a new center and updating other potentials is executed until a predefined condition is reached. This condition depends on the minimum value of the potentials or the required number of clusters which are reached.

This method consists in computing a potential value for each point (θ_i , $i = 1, \dots, N$), based on its distances to the other data points and consider each data point as a potential cluster center. The potential is computed using the following expression:

$$P_i = \sum_{j=1}^N e^{-\frac{4}{r_a} \|\theta_i - \theta_j\|^2}. \quad (9)$$

The potential of each local parameter is a function of the distance from this parameter to all the other local parameters. Thus, a local parameter with many neighboring local parameters will have the highest potential value. The constant r_a is the radius defining the neighborhood which can be determined by the following expression:

$$r_a = \frac{\alpha}{N} \sum_{i=1}^N \frac{1}{n_\rho} \sum_{j=1}^{n_\rho} \|\theta_i - \theta_j\|. \quad (10)$$

where α can be chosen as follows $0 < \alpha < 1$.

Equation (9) can be exploited to eliminate the outliers. As this equation attribute to the outliers a low potential, we can fix a threshold γ under which the local parameters are not accepted and then removed from the data set. This threshold is described by the following equation:

$$\gamma = \min(P) + \beta(\max(P) - \min(P)). \tag{11}$$

where P is the vector containing the potentials P_i such that $P = [P_1, \dots, P_N]$ and β is a parameter chosen as $0 < \beta < 1$.

The elimination of outliers reduces the parameter vectors to $(\theta_i, i = 1, \dots, N')$ ($N' < N$). Then, from this new data set, we select the data point with the highest potential value as the first cluster center.

Let θ_1^* be this first center and P_1^* be its potential. The other potentials P_j , ($j = 1, \dots, N'$) are then updated using this expression:

$$P_i \Leftarrow P_i - P_1^* e^{-\frac{4}{r_b} \|\theta_i - \theta_1^*\|^2}. \tag{12}$$

Expression (13) allows to associate lower potentials to the local parameters close to the first center. Consequently, this choice guaranties that these parameters are not selected as cluster centers in the next step. The parameter r_b is a positive constant that must be chosen larger than r_a to avoid obtaining cluster centers which are too close to each other. The constant r_b is computed using this formula:

$$r_b = \frac{\alpha}{N} \sum_{i=1}^N \max_{j=1:n_p} (\|\theta_i - \theta_j\|). \tag{13}$$

In general after obtaining the k th cluster center, the potential of every local parameter is updated by the following formula:

$$P_i \Leftarrow P_i - P_k^* e^{-\frac{4}{r_b} \|\theta_i - \theta_k^*\|^2}. \tag{14}$$

where P_k^* and θ_k^* are respectively the potential and the center of the k th local parameter.

The number of sub-models s is a parameter that we would like to determine. Therefore, we have developed some criteria for accepting or rejecting the cluster centers as it is explained in the algorithm of the next section.

To search the elements belonging to each cluster, we compute the distance between the estimated output and the real one and classify $\varphi(k)$ within the cluster which has the minimum distance.

$$\arg \min(\theta_i^T \overline{\varphi}_k - y_k), \quad i = 1, \dots, s. \tag{15}$$

The Chiu's clustering technique can be summarized by the following algorithm:

Algorithm 1: Identification algorithm

Data: Dispose of $\{\theta_i\}_{i=1}^N$ from a given data set (φ_i, y_i)

Main steps:

- Compute P_i for every $\{\theta_i\}_{i=1}^N$ according to (9)
- Determine the filtered local parameters $\{\theta_i\}_{i=1}^{N'}$, ($N' < N$)
- Compute the first cluster center θ_1^* from (9)

repeat

Compute the other cluster centers according to the updated potential formula (14)

if $P_k^* > \gamma$ **then**

Compute $V(c)$ such as:

$$V(c) = \|\theta_k^* - \theta_c^*\|, \quad c = 1, \dots, k-1. \quad (16)$$

where θ_k^* is the current cluster center and θ_c^* , $c = 1, \dots, k-1$ are the last selected ones.

if $V(c) > \varepsilon$, $c = 1, \dots, k-1$ **then**

| accept θ_k^* as a cluster center and continue

else

| reject θ_k^* and compute a new potential

end

else

| reject θ_k^* and break

end

until $V(c) \leq \varepsilon$, $c = 1, \dots, k-1$;

Result: Determination of the number of clusters s and the parameters $\{\theta_i\}_{i=1}^s$

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

4.2 The DBSCAN Clustering Technique

The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm is a pioneer algorithm of density-based clustering (Chaitali 2012; Sander et al. 1998). This algorithm is based on the concepts of density-reachability and density-connectivity. These concepts depend on two input parameters: epsilon (ε) and (*MinPts*).

- ε : is the radius around an object that defines its ε -neighborhood.
- *MinPts*: is the minimum number of points.

For a given object q , when the number of objects within the ε -neighborhood is at least *MinPts*, then q is defined as a core object. All objects within its ε -neighborhood are said to be directly density-reachable from q .

In general, an object p is considered density-reachable if it is within the ε -neighborhood of an object that is directly density-reachable or just density-

reachable from q . The objects p and q are said to be density-connected if there exist an object g that both p and q are density-reachable from.

The DBSCAN algorithm define then a cluster as the set of objects in a data set that are density-connected to a particular core object. Any object that is not part of a cluster is categorized as noise. For a given data set $S = \{\theta_k\}_{k=1}^N$, ε and $MinPts$ as inputs, the ε -neighborhood of a point θ_i is defined as:

$$N_\varepsilon(\theta_i) = \{\theta_j \in S; \|\theta_i - \theta_j\| \leq \varepsilon\} \tag{17}$$

The DBSCAN constructs clusters by checking the ε -neighborhood of each object in the data set. If the cardinal of the ε -neighborhood (denoted by cN_ε) of an object θ_k contains more than $MinPts$, a new cluster is created having θ_k as core. The DBSCAN then iteratively collects directly density-reachable objects from these core objects. The process terminates when no new objects can be added to any cluster. The main steps of this algorithm can be summarized as follows:

Algorithm 2: DBSCAN algorithm

Data: Define the input parameters: $S = \{\theta_k\}_{k=1}^N$, $MinPts$ and ε

Main steps:

```

for  $k=1:N$  do
    if  $\theta_k$  is not in a cluster then
        Compute  $N_\varepsilon(\theta_k)$ 
        if  $cN_\varepsilon(\theta_k) < MinPts$  then
            | Mark  $\theta_k$  as noise
        else
            cluster-label=cluster-label+1
            for  $j=1:cN_\varepsilon(\theta_k)$  do
                | Mark all point in  $N_\varepsilon(\theta_k)$  with the current cluster label
            end
            Lend  $N_\varepsilon(\theta_k)$  to the Seed list  $L_S = [L_S \ N_\varepsilon(\theta_k)]$ 
            while  $L_S$  is not empty do
                 $\theta_r = L_S(1)$ 
                Compute  $N_\varepsilon(\theta_r)$ 
                if  $cN_\varepsilon(\theta_r) \geq MinPts$  then
                    for  $o=1:cN_\varepsilon(\theta_r)$  do
                        if  $\theta_o$  is not in a cluster nor marked as noise then
                            Mark  $\theta_o$  with the current cluster-label
                            Lend  $N_\varepsilon(\theta_r)$  to the Seed list  $L_S = [L_S \ N_\varepsilon(\theta_r)]$ 
                             $L_S(1) = [ \ ]$ 
                        end
                    end
                else
                    |  $L_S(1) = [ \ ]$ 
                end
            end
        end
    end
end
end

```

Result: Determination of the number of clusters s and the parameters $\{\theta_i\}_{i=1}^s$

5 Simulation Examples

In this section, we aim at illustrating the performance of the proposed methods with three simulation examples. First of all, we take an academic PWARX model where the proposed methods are compared with the well known k-means one (Ferrari-Trecate et al. 2001, 2003). After that, a nonlinear model is considered to show the efficiency of the proposed methods in approximating any nonlinear model. Finally, a pH neutralization process is simulated in order to prove their ability to model complex systems and to determine the number of sub-models.

5.1 Quality Measures

To achieve the purpose of these simulations, we consider the following quality measures (Juloski et al. 2006):

- The maximum of relative error of parameter vectors is defined by

$$\Delta_\theta = \max_{i=1,\dots,s} \frac{\|\theta_i - \bar{\theta}_i\|_2}{\|\bar{\theta}_i\|_2} \quad (18)$$

where $\bar{\theta}_i$ and θ_i are the true and the estimated parameter vectors for sub-model i , respectively. The identified model is deemed acceptable if Δ_θ is small or close to zero.

- The averaged sum of the squared residuals is defined by

$$\sigma_e^2 = \frac{1}{s} \sum_{i=1}^s \frac{SSR_i}{|D_i|} \quad (19)$$

where $SSR_i = \sum_{(y(k), \varphi(k)) \in D_i} (y(k) - [\varphi(k)' 1] \theta_i)^2$ and $|D_i|$ is the cardinality of cluster D_i .

The identified model is considered acceptable if σ_e^2 is small and/or close to the expected noise variance of the true system.

- The percentage of the output variation that is explained by the model is defined by

$$FIT = 100 \cdot \left(1 - \frac{\|\hat{y} - y\|_2}{\|y - \bar{y}\|_2} \right) \quad (20)$$

where \hat{y} and y are the estimated and the real outputs' vectors, respectively, and \bar{y} is the mean value of y .

The identified model is considered acceptable if FIT is close to 100.

- The relative error expressed in percentage (%) is given by:

$$e_r(k) = 100 \cdot \frac{|y(k) - \hat{y}(k)|}{|y(k)|} \tag{21}$$

where $\hat{y}(k)$ and $y(k)$ are the estimated and the real outputs at time k . The identified model is considered acceptable if e_r is close to 0 %.

5.2 Identification Results of a PWARX Model

Consider the following PWARX model (Boukharouba 2011):

$$y(k) = \begin{cases} [0.4 & 0.5 & 0.3] \bar{\varphi}(k) + e(k) & \text{if } \varphi(k) \in H_1, \\ [-0.7 & 0.6 & -0.5] \bar{\varphi}(k) + e(k) & \text{if } \varphi(k) \in H_2, \\ [0.4 & -0.2 & -0.2] \bar{\varphi}(k) + e(k) & \text{if } \varphi(k) \in H_3, \end{cases} \tag{22}$$

$$\begin{aligned} H_1 &= \{ \varphi \in \mathbb{R}^2 : [1 \quad 0.3 \quad 0] \bar{\varphi} \geq 0 \text{ and } [0 \quad 0.5 \quad 0] \bar{\varphi} > 0 \} \\ H_2 &= \{ \varphi \in \mathbb{R}^2 : [1 \quad 0.3 \quad 0] \bar{\varphi} \leq 0 \text{ and } [1 \quad -0.3 \quad 0] \bar{\varphi} < 0 \} \\ H_3 &= \{ \varphi \in \mathbb{R}^2 : [1 \quad -0.3 \quad 0] \bar{\varphi} \geq 0 \text{ and } [0 \quad 0.5 \quad 0] \bar{\varphi} \leq 0 \} \end{aligned} \tag{23}$$

where $s = 3, n_a = 1, n_b = 1$, and $\varphi(k) = [y(k - 1) \quad u(k - 1)]^T$ is the regressor vector.

System (22) is simulated using an input signal $u(k)$ and a noise signal $e(k)$ which are normal distributions with variances respectively 0.5 and 0.05. The output $y(k)$ is presented in Fig. 1.

Table 1 presents the estimated parameter vectors obtained with the proposed methods and the k-means one.

Fig. 1 The real output of the system (*squares*: output of sub-model 1, *triangles*: output of sub-model 2 and *dots*: output of sub-model 3)

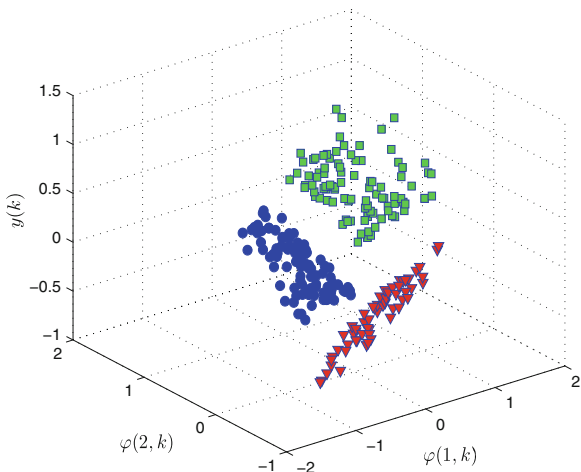


Table 1 Estimated parameters

	True values	Chiu ($n_p = 20$)	DBSCAN ($n_p = 21$)	k-means ($n_p = 7$)
θ_1	$\begin{bmatrix} 0.4 \\ 0.5 \\ 0.3 \end{bmatrix}$	$\begin{bmatrix} 0.4046 \\ 0.5138 \\ 0.2919 \end{bmatrix}$	$\begin{bmatrix} 0.4054 \\ 0.4903 \\ 0.2992 \end{bmatrix}$	$\begin{bmatrix} 0.4064 \\ 0.5464 \\ 0.2598 \end{bmatrix}$
θ_2	$\begin{bmatrix} -0.7 \\ 0.6 \\ -0.5 \end{bmatrix}$	$\begin{bmatrix} -0.6179 \\ 0.5336 \\ -0.4740 \end{bmatrix}$	$\begin{bmatrix} -0.7369 \\ 0.6675 \\ -0.5239 \end{bmatrix}$	$\begin{bmatrix} -0.6955 \\ 0.5903 \\ -0.4939 \end{bmatrix}$
θ_3	$\begin{bmatrix} 0.4 \\ -0.2 \\ -0.2 \end{bmatrix}$	$\begin{bmatrix} 0.4015 \\ -0.2071 \\ -0.2042 \end{bmatrix}$	$\begin{bmatrix} 0.4679 \\ -0.1977 \\ -0.2298 \end{bmatrix}$	$\begin{bmatrix} 0.4792 \\ -0.2101 \\ -0.2406 \end{bmatrix}$

After obtaining the estimated parameter vectors, we apply the SVM algorithm in order to estimate the regions. We can then attribute each parameter vector to the corresponding region where it is defined. The estimated outputs obtained with three algorithms are presented in Fig. 2.

Table 2 presents the quality measures (18), (19) and (20) of the two proposed methods and the k-means method. The obtained results prove the efficiency of the proposed methods compared with the existing method (k-means).

5.3 Identification Results of a Nonlinear Model

Consider the nonlinear system described by the following equation (Lai et al. 2010):

$$y(k) = \frac{1.5y(k-1)y(k-2)}{1 + y^2(k-1) + y^2(k-2)} + \sin(y(k-1) + y(k-2)) + u(k-1) + 0.8u(k-2) \tag{24}$$

This nonlinear system can be modeled by a PWARX model of the form (Lai 2011):

$$y(k) = \begin{cases} \theta_1^T \bar{\varphi}(k) & \text{if } \varphi \in H_1 \\ \vdots \\ \theta_s^T \bar{\varphi}(k) & \text{if } \varphi \in H_s \end{cases} \tag{25}$$

where

$$\varphi(k) = [y(k-1), y(k-2), u(k-1), u(k-2)]^T \tag{26}$$

$$\bar{\varphi} = [\varphi^T \quad 1]^T. \tag{27}$$

θ_i are the parameter vectors and s is the number of submodels to be determined. $u(k)$ is a random input in the range of $[-2, 2]$.

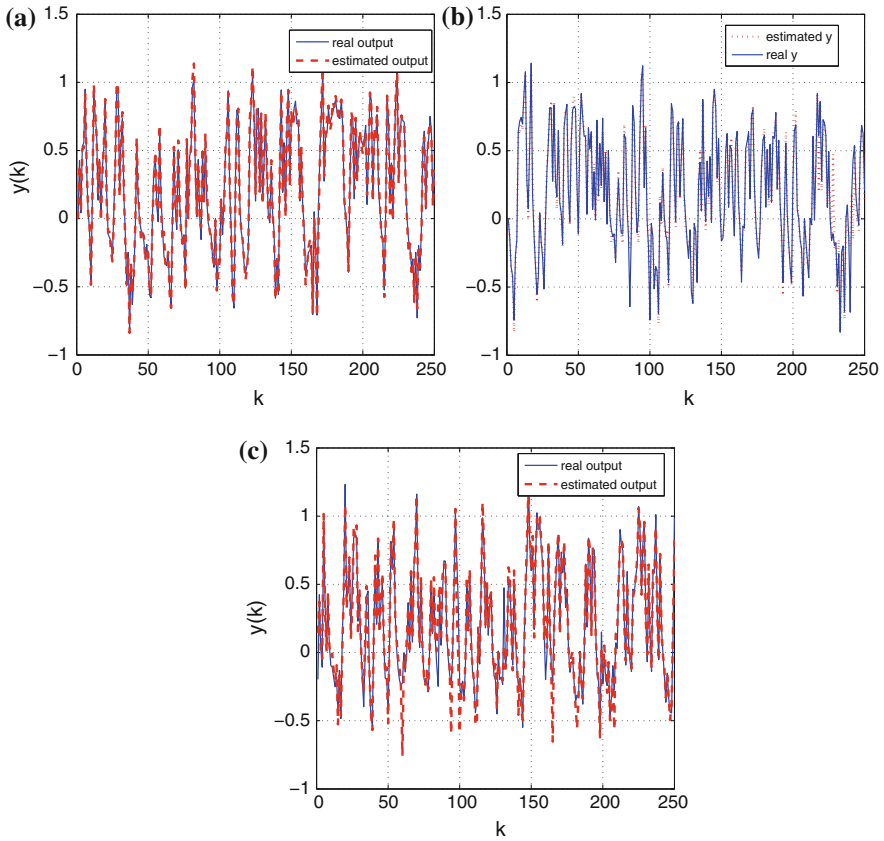


Fig. 2 The estimated outputs **a** with Chiu, **b** with DBSCAN, and **c** with k-means

Table 2 Validation results

Quality measures	Chiu	DBSCAN	k-means
$\Delta\theta$	0.0876	0.1514	0.1828
σ_e^2	0.0023	0.0097	0.0109
<i>FIT</i>	89.4191	79.0183	74.195

For the DBSCAN based method, the choice of the synthesis parameters n_ρ , *MinPts* and ϵ is as follows:

$$\begin{cases} n_\rho = 20; \\ \textit{MinPts} = 35; \\ \epsilon = 0.85 \end{cases}$$

For the Chiu clustering algorithm, we have only one synthesis parameter: $n_\rho = 17$.

The number of submodels s depends on the initial parameters chosen. With the parameters described above, we obtain $s = 6$.

The parameter vectors are presented in Tables 3 and 4.

Figures 3 and 4 illustrate the outputs and the relative error signals of the two proposed methods.

In Table 5, the *FIT* is computed for the identification and the validation with the two proposed methods. The obtained results are very satisfactory and show that the performances of the two methods are close.

5.4 Identification Results of a PH Neutralisation Process

5.4.1 Process Description

The ‘neutralization’ is used to describe the reaction result between an acid and a base in which the properties of H^+ and OH^- that characterized the acid and base will be destroyed or neutralized. In fact, the ions H^+ and OH^- will be combined to form the water molecule H_2O . The resulting solution produced by the reaction is

Table 3 Estimated parameter vectors with the Chiu’s method

Parameters vector	Estimated values with Chiu
θ_1	$[1.0996 \ 0.6338 \ 1.0259 \ 0.8247 \ 0.0762]^T$
θ_2	$[-0.7769 \ -1.0611 \ 1.0016 \ 0.8446 \ -2.2539]^T$
θ_3	$[0.5066 \ 0.9219 \ 0.9781 \ 0.7686 \ 0.1886]^T$
θ_4	$[-0.6576 \ -0.4256 \ 0.9388 \ 0.8097 \ -1.4943]^T$
θ_5	$[-0.3013 \ -0.5044 \ 1.0644 \ 0.6904 \ 1.7459]^T$
θ_6	$[-0.8306 \ -0.5365 \ 1.0030 \ 0.7557 \ 2.7559]^T$

Table 4 Estimated parameter vectors with the DBSCAN method

Parameters vector	Estimated values with DBSCAN
θ_1	$[0.5134 \ 0.4871 \ 1.0377 \ 0.8082 \ 0.0187]^T$
θ_2	$[-0.5974 \ -0.3021 \ 0.8905 \ 0.7355 \ 2.1231]^T$
θ_3	$[0.7044 \ 0.0158 \ 1.0153 \ 0.8255 \ 0.1468]^T$
θ_4	$[-0.6958 \ -0.6557 \ 1.0714 \ 0.7435 \ -1.5669]^T$
θ_5	$[0.3111 \ 0.0277 \ 0.9738 \ 0.7616 \ -0.3855]^T$
θ_6	$[-0.6177 \ -0.3720 \ 0.9606 \ 0.7291 \ -1.2358]^T$

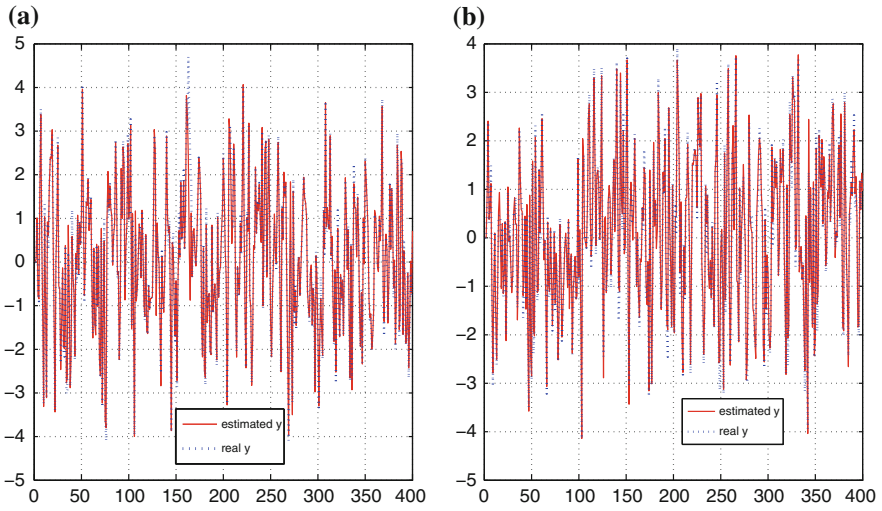


Fig. 3 The estimated outputs **a** with Chiu, and **b** with DBSCAN

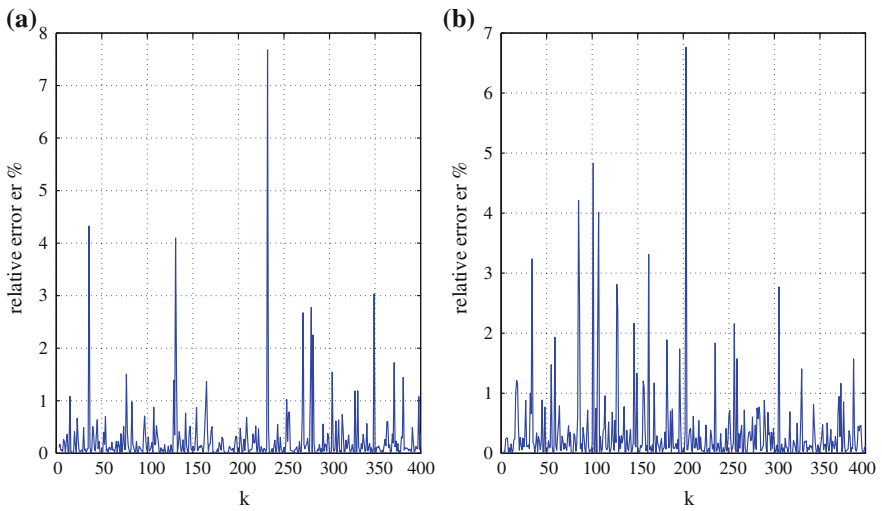


Fig. 4 The relative error **a** with Chiu, and **b** with DBSCAN

Table 5 Quality measures with the two proposed methods

	Chiu	DBSCAN
<i>FIT</i> for identification	78.0543	84.1135
<i>FIT</i> for validation	75.2050	74.1906

composed of a salt and water. The general formula for acid–base neutralization reactions can be written as:



The process of pH neutralization (see Fig. 5) is constituted essentially of a treatment tank of cross sectional area A , a mixer, acid and base injection pipes, a pH probe, a level sensor to measure the level h in the tank and a discharge valve (Henson and Seborg 1994; Salehi et al. 2009). It consists of an acid stream q_1 , buffer stream q_2 and base stream q_3 that are mixed in the tank. The effluent stream q_4 exits the tank via the discharge valve with an adjusted pH_m . The streams $\{q_i\}_{i=1}^4$ are characterized by the following parameters:

- $\{W_{ai}\}_{i=1}^4$ are the charge related quantities for $\{q_i\}_{i=1}^4$.
- $\{W_{bi}\}_{i=1}^4$ are the mass balance quantities for $\{q_i\}_{i=1}^4$.

The pH probe introduces a delay time τ in the measured pH_m value such as $pH_m = pH(t - \tau)$.

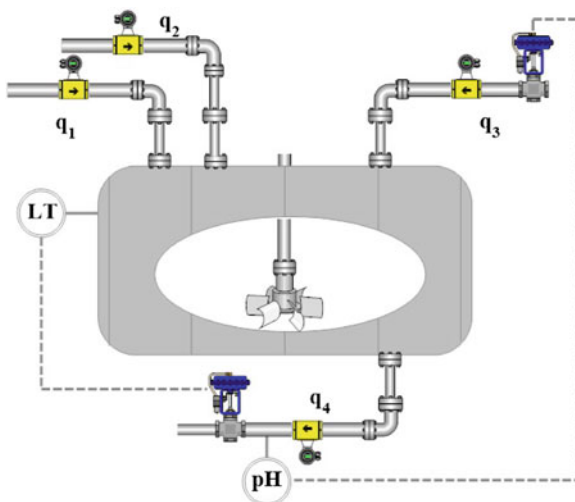
The objective of the pH neutralization process is to control the pH value of the effluent through manipulating the base flow rate q_3 while considering the acid flow rate q_1 and the buffer flow rate q_2 as disturbances.

The dynamic model of the neutralization process is developed as follows:

- The pH value of the obtained solution is derived from the conservation equations and equilibrium reactions as follows:

$$W_{a4} + \frac{K_w}{[H^+]} + W_{b4} \frac{\frac{K_{a1}}{[H^+]} + \frac{2K_{a1}K_{a2}}{[H^+]^2}}{1 + \frac{K_{a1}}{[H^+]} + \frac{K_{a1}K_{a2}}{[H^+]^2}} - [H^+] = 0. \tag{29}$$

Fig. 5 A pH neutralization process



Knowing that

$$pH_m = -\log([H^+]) \quad (30)$$

$$K_w = [H^+][OH^-], \quad (31)$$

Equation (29) can be then rewritten as:

$$W_{a4} + 10^{pH_m-14} + W_{b4} \frac{1 + 2(10^{pH_m-pK_{a2}})}{1 + 10^{pH_m-pK_{a1}} + 10^{pH_m-pK_{a2}}} - 10^{-pH_m} = 0 \quad (32)$$

- The mass balance yields to:

$$A \frac{dh}{dt} = q_1 + q_2 + q_3 - q_4 \quad (33)$$

Taking into account that the exit flow rate $q_4 = C_v \cdot h^{0.5}$, Eq. (33) becomes:

$$A \frac{dh}{dt} = q_1 + q_2 + q_3 - C_v \cdot h^{0.5} \quad (34)$$

where C_v is the constant valve coefficient.

- The differential equations of the effluent reaction invariants (W_{a4} , W_{b4}) can be determined as follows:

$$A h \frac{dW_{a4}}{dt} = q_1(W_{a1} - W_{a4}) + q_2(W_{a2} - W_{a4}) + q_3(W_{a3} - W_{a4}) \quad (35)$$

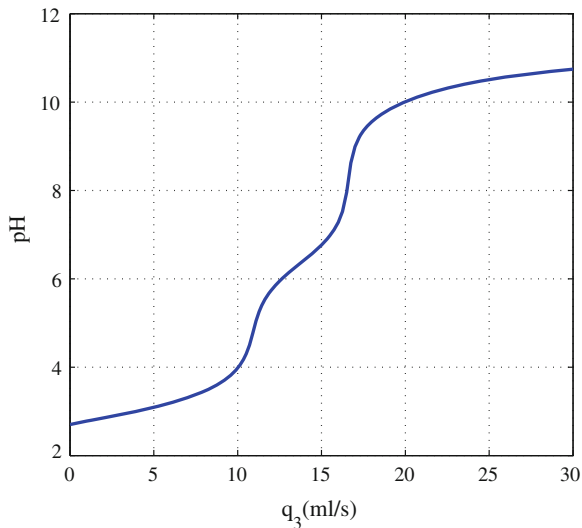
$$A h \frac{dW_{b4}}{dt} = q_1(W_{b1} - W_{b4}) + q_2(W_{b2} - W_{b4}) + q_3(W_{b3} - W_{b4}) \quad (36)$$

Nominal model parameters and operating conditions (Xiao et al. 2014) are given in Table 6.

The static nonlinearity of this process can be represented by the titration curve shown in Fig. 6 with a beginning pH of 2.7 and an ending pH of 10.7. A brief glance at the curve indicates that the process of pH neutralization is highly nonlinear.

Table 6 Operation parameters of the pH neutralization process

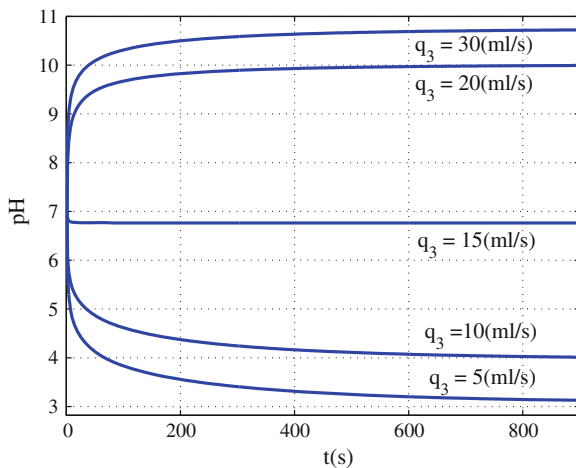
$q_1 = 16.6$ ml/s	$W_{a1} = 3 \times 10^{-3}$ mol/l
$q_2 = 0.55$ ml/s	$W_{a2} = -3 \times 10^{-2}$ mol/l
$q_3 = 15.6$ ml/s	$W_{a3} = -3.05 \times 10^{-3}$ mol/l
$h = 14.0$ cm	$W_{b1} = 0$
$A = 207$ cm ²	$W_{b2} = 3 \times 10^{-2}$ mol/l
$C_v = 8.75$ ml/cm/s	$W_{b3} = 5 \times 10^{-5}$ mol/l
$pK_{a1} = 6.35$	$pH_4 = 7$
$pK_{a2} = 10.25$	$\tau = 0.5$

Fig. 6 The titration curve

5.4.2 Structure Identification

It was mentioned that the early approaches of identification of pH neutralization process approximate this process around an operating range as a First Order Plus Delay Time model. Added to that, the evolution of the pH in Fig. 7, for a fixed values of the input q_3 , is similar to a first order system response.

Therefore, we propose to represent the sub-models by a discrete first order plus dead time models ($n_a = 1$, $n_b = 2$) defined by:

Fig. 7 The pH evolution with different values of q_3 

$$y(k) = \begin{cases} a_{1,1}y(k-1) + b_{1,1}u(k-1) + b_{1,2}u(k-2) \\ \quad \text{if } \varphi(k) \in H_1 \\ \vdots \\ a_{s,1}y(k-1) + b_{s,1}u(k-1) + b_{s,2}u(k-2) \\ \quad \text{if } \varphi(k) \in H_s \end{cases} \quad (37)$$

where the regressor vector is defined by:

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

and the parameter vectors are denoted by:

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

5.4.3 Input Design

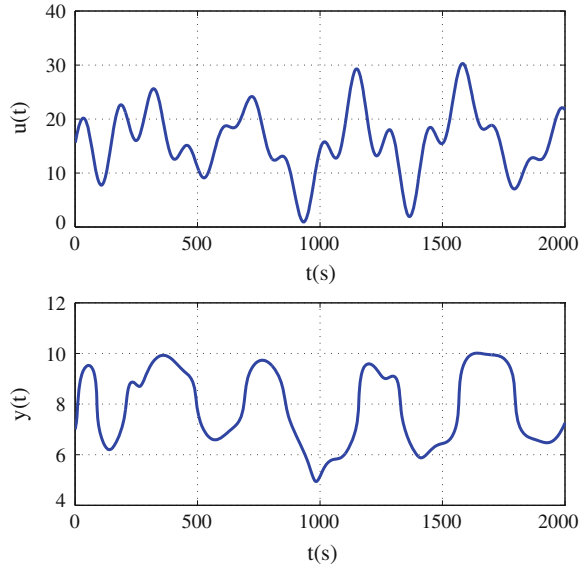
The input design is an important aspect to be considered when implementing nonlinear system identification experiments. In fact, two main properties must be verified by this input in order to generate representative data measurements to be used in identification purpose. First, the input must be able to excite the totality of dynamics range present in the system. Second, the used input signal must illustrate the response of the system to a range of amplitude changes since these models have nonlinear gains. For these reasons, we have considered the Multi-Sine sequence as input sequences to identify the pH neutralization process since it satisfies the above two conditions. It presents several frequencies and exhibits different amplitude changes. The dynamic of this input is defined according to the dominant time constant range of the process. The amplitudes are selected to cover the totality operating region around the nominal value of the base flow rate $q_3 = 15.6$ ml/s.

5.4.4 Results

The nonlinear model of the pH process defined by Eqs. (32), (34), (35) and (36) and the parameters of Table 6 is used to generate the output using a Multi-Sine excitation Sequence. The system output is corrupted by a Gaussian white noise with zero mean and standard deviation $\sigma = 0.001$ in order to simulate industrial situations where the obtained measurements are often noisy. The obtained input-output data illustrated in Fig. 8 are then divided into two parts. The first part is used for the identification and the second is considered for the validation purpose.

The number of neighboring is chosen $n_\rho = 85$ for the two methods. The DBSCAN approach uses the following synthesis parameters

Fig. 8 The data of the multi-sine input



$$\begin{cases} \text{MinPts} = 40; \\ \varepsilon = 0.18 \end{cases}$$

The number of submodels obtained with these parameters is ($s = 6$). The parameter vectors are illustrated in Table 7.

The validation results and the estimated titration curves are presented respectively in Figs. 9 and 10 which shows that the obtained model gives good results in terms of dynamic and nonlinear gain of the pH process.

Now, we compare the performance of the two proposed methods using the quality measures (19), (20) and (21). The obtained results are summarized in Table 8 and Fig. 11.

Table 7 Estimated parameter vectors

Parameter vectors	Estimated values DBSCAN	Estimated values Chiu
θ_1	$[0.9906 \quad 0.0264 \quad -0.0218]^T$	$[0.9855 \quad 0.0179 \quad -0.0101]^T$
θ_2	$[0.9808 \quad 0.0287 \quad -0.0190]^T$	$[0.9692 \quad 0.0593 \quad -0.0423]^T$
θ_3	$[0.9763 \quad 0.0282 \quad -0.0152]^T$	$[1.0102 \quad -0.0216 \quad 0.0200]^T$
θ_4	$[1.0272 \quad 0.0343 \quad -0.0434]^T$	$[1.0351 \quad 0.0273 \quad -0.0550]^T$
θ_5	$[0.9968 \quad -0.0390 \quad 0.0432]^T$	$[0.9632 \quad -0.0289 \quad 0.0440]^T$
θ_6	$[0.9690 \quad -0.0404 \quad 0.0533]^T$	$[1.0021 \quad 0.0903 \quad -0.0910]^T$

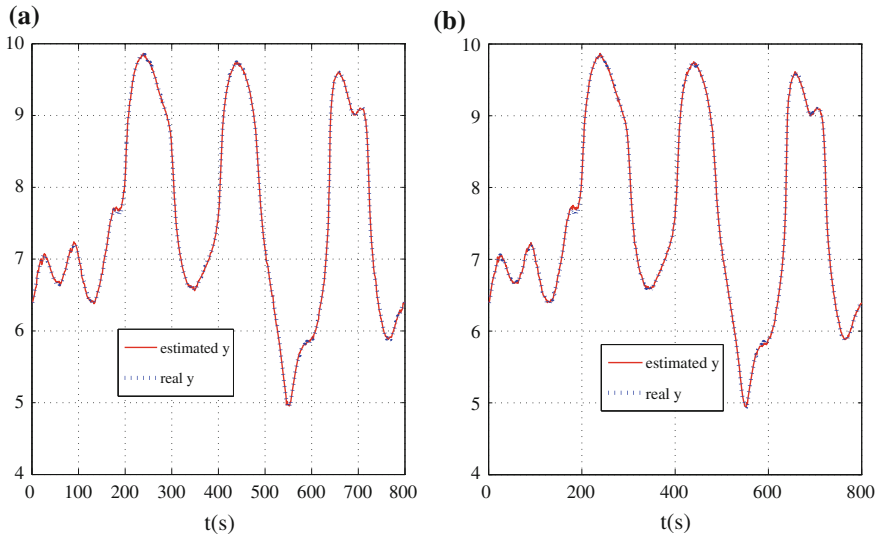


Fig. 9 The validation outputs **a** with Chiu, and **b** with DBSCAN

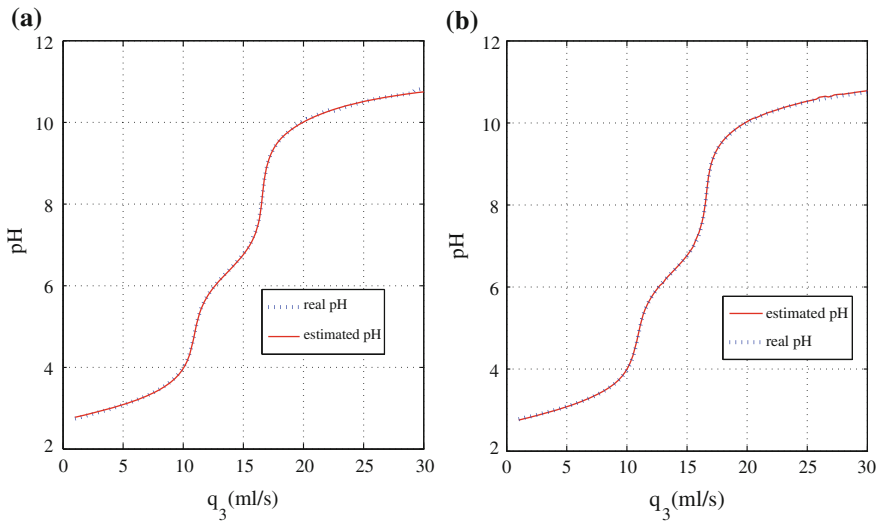


Fig. 10 The validation of the titration curve **a** with Chiu, and **b** with DBSCAN

Table 8 Quality measures with the two proposed methods

	Chiu	DBSCAN
<i>FIT</i> for identification	98.0203	97.8688
<i>FIT</i> for validation	98.1254	98.0405
σ_e^2	8.8552×10^{-4}	0.0013

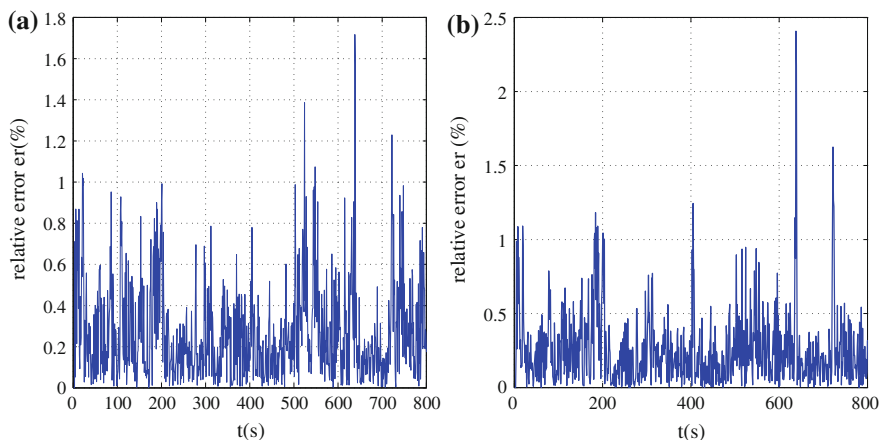
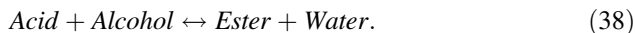


Fig. 11 Relative error for the two proposed methods **a** with Chiu, and **b** with DBSCAN

6 Experimental Example: A Semi-batch Reactor

6.1 Process Description

The olive oil esterification reactor produces 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:



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 removed 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 block diagram of the process is shown in Fig. 12. It 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-

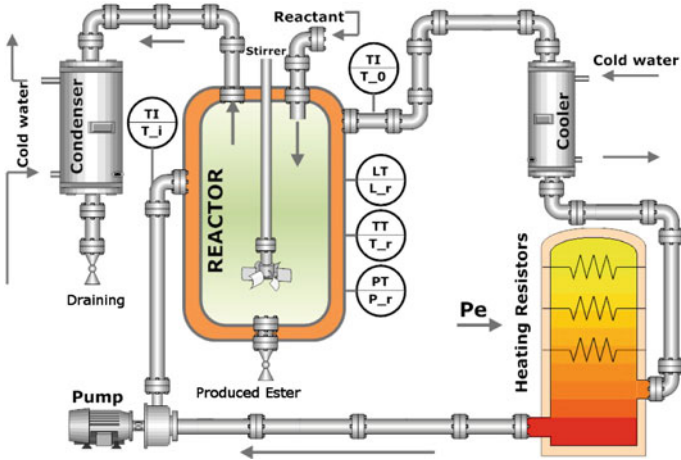


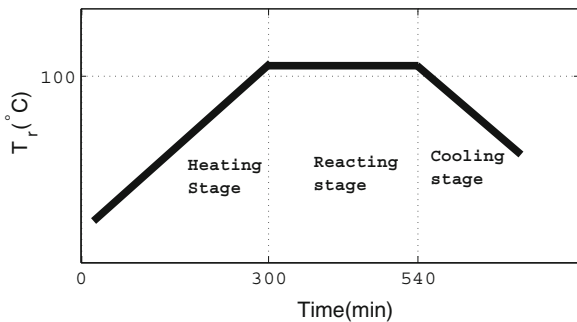
Fig. 12 Block diagram of the reactor

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 ester production by this reactor is based on three main steps as illustrated in Fig. 13.

Fig. 13 Specific trajectory of the reactor temperature



6.2 Experimental Results

The alternative of considering a PWA map is very interesting because the characteristic of the system can be considered as piecewise linear in each operating phase: the heating phase, the reacting phase and the cooling phase.

Previous works has demonstrated that the adequate estimated orders n_a and n_b of each sub-model are equal to two (Talmoudi et al. 2008). Thus, we can adopt the following structure:

$$y(k) = \begin{cases} -a_{1,1}y(k-1) - a_{1,2}y(k-2) + b_{1,1}u(k-1) \\ + b_{1,2}u(k-2) & \text{if } \varphi(k) \in H_1 \\ \vdots \\ a_{s,1}y(k-1) + a_{s,2}y(k-2) + b_{s,1}u(k-1) \\ + b_{s,2}u(k-2) & \text{if } \varphi(k) \in H_s \end{cases} \quad (39)$$

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 is denoted by:

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

We have picked out some input-output measurements from the reactor in order to identify a model to this process. We have taken two measurement files, one for the identification having a length $N = 220$ and another one of length $N = 160$ for the validation.

The measurement file used in this identification is presented in Fig. 14.

Fig. 14 The real input-output evolution

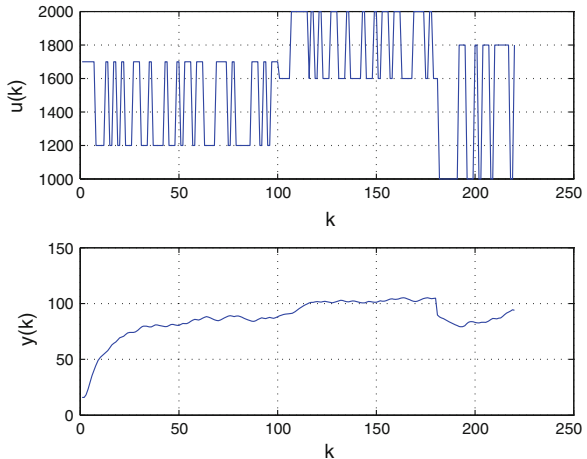


Table 9 Estimated parameter vectors with the proposed clustering techniques

Parameter vectors	Estimated parameters with Chiu	Estimated parameters with DBSCAN
θ_1	$\begin{bmatrix} -1.4256 \\ 0.4508 \\ 4.9853 \times 10^{-4} \\ 0.0010 \end{bmatrix}$	$\begin{bmatrix} -1.4404 \\ 0.4692 \\ 0.0003 \\ 0.0014 \end{bmatrix}$
θ_2	$\begin{bmatrix} -1.1604 \\ 0.2111 \\ 0.0015 \\ 0.0014 \end{bmatrix}$	$\begin{bmatrix} -1.1144 \\ 0.1772 \\ 0.0003 \\ 0.0032 \end{bmatrix}$
θ_3	$\begin{bmatrix} -1.0847 \\ 0.1490 \\ -3.9782 \times 10^{-4} \\ 0.0040 \end{bmatrix}$	$\begin{bmatrix} -1.0591 \\ 0.1304 \\ 0.0006 \\ 0.0034 \end{bmatrix}$

We apply the proposed identification procedures in order to represent the reactor by a PWARX model. The number of neighboring is chosen $n_p = 70$ with the two proposed techniques. Our purpose is to estimate the number of sub-models s , the parameter vectors $\theta_i(k)$, $i = 1, \dots, s$ 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 9.

The attribution of every parameter vector to the submodel that has generated it is ensured by the SVM algorithm. The obtained outputs are then computed and they are represented in Fig. 15.

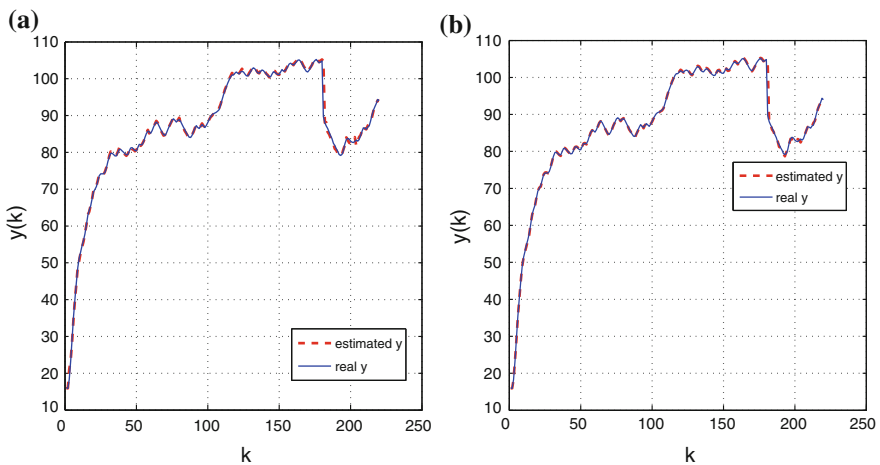


Fig. 15 Estimated outputs with two methods **a** with Chiu, and **b** with DBSCAN

To validate the obtained models, we have considered a new input-output measurement file having a length $N = 160$ shown in Fig. 16.

The real and the estimated validation outputs and the errors are presented in Fig. 17.

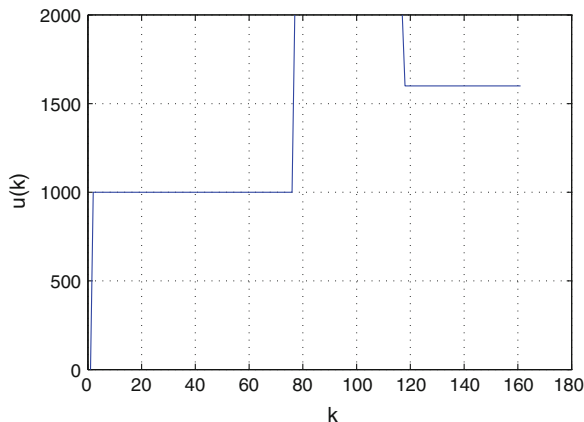


Fig. 16 The validation input

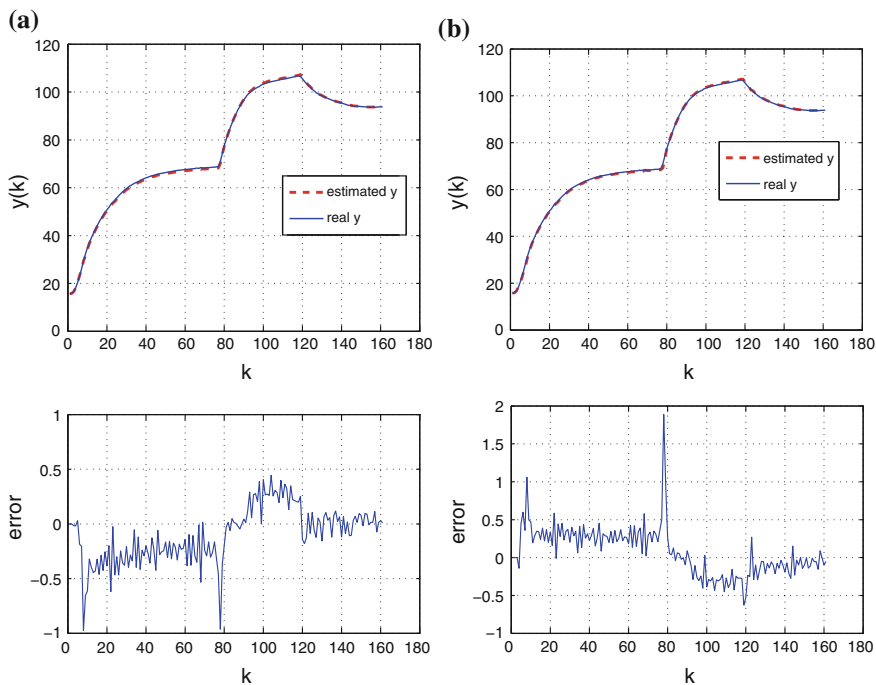


Fig. 17 Estimated validation outputs and the errors with two methods **a** with Chiu, and **b** with DBSCAN

7 Conclusion

In this chapter, we have considered only the clustering based procedures for the identification of PWARX systems. We focused on the most challenging step which is the task of data points classification. In fact, we have proposed the use of two clustering techniques which are the Chiu's clustering algorithm and the DBSCAN algorithm. These algorithms present several advantages. Firstly, they do not require any initialization so the problem of convergence towards local minima is overcome. Secondly, these algorithms are able to remove the outliers from the data set. Finally, our approaches generate automatically the number of sub-models. Numerical simulation results are presented to demonstrate the performance of the proposed approaches and to compare them with the k-means one. Also, an experimental validation with an olive oil reactor is presented to illustrate the efficiency of the developed methods.

References

- Bako, L. (2011). Identification of switched linear systems via sparse optimization. *Automatica*, 47(4), 668–677.
- Bako, L., & Lecoche, S. (2013). A sparse optimization approach to state observer design for switched linear systems. *Systems and Control Letters*, 62(2), 143–151.
- Bemporad, A., Ferrari-Trecate, G., & Morari, M. (2000). Observability and controllability of piecewise affine and hybrid systems. *IEEE Transactions on Automatic Control*, 45(10), 1864–1876.
- Bemporad, A., Garulli, A., Paoletti, S., & Vicino, A. (2003). A greedy approach to identification of piecewise affine models. In *Hybrid systems: Computation and control* (pp. 97–112). New York: Springer.
- Bemporad, A., Garulli, A., Paoletti, S., & Vicino, A. (2005). A bounded-error approach to piecewise affine system identification. *IEEE Transactions on Automatic Control*, 50(10), 1567–1580.
- Boukharouba, K. (2011). Modélisation et classification de comportements dynamiques des systèmes hybrides. *Ph.D. thesis*, Université de Lille, France.
- Chaitali, C. (2012). Optimizing clustering technique based on partitioning DBSCAN and ant clustering algorithm. *International Journal of Engineering and Advanced Technology (IJEAT)*, 2(2), 212–215.
- Chiu, S. (1994). Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems*, 2(3), 267–278.
- Chiu, S. (1997). Extracting fuzzy rules from data for function approximation and pattern classification. In D. Dubois, et al. (Eds.), *Chapter 9 in fuzzy information engineering: A guided tour of applications*. New York: Wiley.
- De Schutter, B., & De Moor, B. (1999). The extended linear complementarity problem and the modeling and analysis of hybrid systems. In *Hybrid systems V* (pp. 70–85). New York: Springer.
- De Schutter, B., & Van den Boom, T. (2000). On model predictive control for max-min-plus-scaling discrete event systems. Technical report, bds 00-04: Control Systems Engineering, Faculty of Information Technology and Systems, Delft University of Technology, The Netherlands.

- Doucet, A., Gordon, N., & Krishnamurthy, V. (2001). Particle filters for state estimation of jump markov linear systems. *IEEE Transactions on Signal Processing*, 49(3), 613–624.
- Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification* (2nd ed.). New York: Wiley.
- Ferrari-Trecate, G., Muselli, M., Liberati, D., & Morari, M. (2001). A clustering technique for the identification of piecewise affine systems. In *Hybrid systems: Computation and control* (pp. 218–231). New York: Springer.
- Ferrari-Trecate, G., Muselli, M., Liberati, D., & Morari, M. (2003). A clustering technique for the identification of piecewise affine systems. *Automatica*, 39(2), 205–217.
- Heemels, W. P., De Schutter, B., & Bemporad, A. (2001). Equivalence of hybrid dynamical models. *Automatica*, 37(7), 1085–1091.
- Henson, M. A., & Seborg, D. E. (1994). Adaptive nonlinear control of a ph neutralization process. *IEEE Transactions on Control Systems Technology*, 2(3), 169–182.
- Juloski, A., Weiland, S., & Heemels, W. (2005). A bayesian approach to identification of hybrid systems. *IEEE Transactions on Automatic Control*, 50(10), 1520–1533.
- Juloski, A. L., Paoletti, S., & Roll, J. (2006). Recent techniques for the identification of piecewise affine and hybrid systems. In *Current trends in nonlinear systems and control* (pp. 79–99). New York: Springer.
- Lai, C. Y. (2011). Identification and control of nonlinear systems using multiple models. *Ph.D. thesis*.
- Lai, C. Y., Xiang, C., & Lee, T. H. (2010). Identification and control of nonlinear systems via piecewise affine approximation. In *The 49th IEEE Conference on Decision and Control (CDC)* (pp. 6395–6402).
- Lassoued, Z., & Abderrahim, K. (2013a). A comparison study of some PWARX system identification methods. In *The 17th IEEE International Conference in System Theory, Control and Computing (ICSTCC)* (pp. 291–296).
- Lassoued, Z., & Abderrahim, K. (2013b). A Kohonen neural network based method for PWARX identification. In *Adaptation and learning in control and signal processing, IFAC* (Vol. 11, pp. 742–747).
- Lassoued, Z., & Abderrahim, K. (2013c). New approaches to identification of PWARX systems. *Mathematical Problems in Engineering*. <http://dx.doi.org/10.1155/2013/845826>.
- Lassoued, Z., & Abderrahim, K. (2013d). A new clustering technique for the identification of PWARX hybrid models. In *The 9th IEEE Asian Control Conference (ASCC)* (pp. 1–6).
- Lassoued, Z., & Abderrahim, K. (2014a). An experimental validation of a novel clustering approach to PWARX identification. *Engineering Applications of Artificial Intelligence*, 28, 201–209.
- Lassoued, Z., & Abderrahim, K. (2014b). New results on PWARX model identification based on clustering approach. *International Journal of Automation and Computing*, 11(2), 180–188.
- Lin, J., & Unbehauen, R. (1992). Canonical piecewise-linear approximations. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 39(8), 697–699.
- Nakada, H., Takaba, K., & Katayama, T. (2005). Identification of piecewise affine systems based on statistical clustering technique. *Automatica*, 41(5), 905–913.
- Roll, J., Bemporad, A., & Ljung, L. (2004). Identification of piecewise affine systems via mixed-integer programming. *Automatica*, 40(1), 37–50.
- Salehi, S., Shahrokhi, M., & Nejati, A. (2009). Adaptive nonlinear control of ph neutralization processes using fuzzy approximators. *Control Engineering Practice*, 17(11), 1329–1337.
- Sander, J., Ester, M., Kriegel, H.-P., & Xu, X. (1998). Density-based clustering in spatial databases: The algorithm gdbscan and its applications. *Data Mining and Knowledge Discovery*, 2(2), 169–194.
- Talmoudi, S., Abderrahim, K., Abdennour, R. B., & Ksouri, M. (2008). Multimodel approach using neural networks for complex systems modeling and identification. *Nonlinear Dynamics and Systems Theory*, 8(3), 299–316.
- Vander-Schaft, A. J., & Schumacher, J. M. (1998). Complementarity modeling of hybrid systems. *IEEE Transactions on Automatic Control*, 43(4), 483–490.

- Vidal, R., Chiuso, A., & Soatto, S. (2002). Observability and identifiability of jump linear systems. In *Proceedings of the 41st IEEE Conference on Decision and Control* (Vol. 4, pp. 3614–3619).
- Wang, L. (2005). *Support vector machines: Theory and applications* (Vol. 177). New York: Springer.
- Wen, C., Wang, S., Jin, X., & Ma, X. (2007). Identification of dynamic systems using piecewise-affine basis function models. *Automatica*, 43(10), 1824–1831.
- Xiao, C., Xue, A., Peng, D., & Guo, Y. (2014). Modeling of ph neutralization process using fuzzy recurrent neural network and dna based nsga-ii. *Journal of the Franklin Institute*, 351(7), 3847–3864.
- Xu, J., Huang, X., Mu, X., & Wang, S. (2012). Model predictive control based on adaptive hinging hyperplanes model. *Journal of Process Control, Elsevier*, 22(10), 1821–1831.