Delft University of Technology

Delft Center for Systems and Control

Technical report 11-009

A new method for hybrid-fuzzy identification*

A. Núñez, D. Sáez, I. Škrjanc, and B. De Schutter

If you want to cite this report, please use the following reference instead:

A. Núñez, D. Sáez, I. Škrjanc, and B. De Schutter, "A new method for hybrid-fuzzy identification," *Proceedings of the 18th IFAC World Congress*, Milan, Italy, pp. 15013–15018, Aug.–Sept. 2011.

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.24.73 (secretary)
URL: https://www.dcsc.tudelft.nl

A new method for hybrid-fuzzy identification ⋆

Alfredo Núñez * **Doris Sáez** ** **Igor Škrjanc** ***
Bart De Schutter *

* *Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands (e-mail: {a.a.nunezvicencio, b.deschutter}@tudelft.nl).*
** *Electrical Engineering Department, Universidad de Chile, Santiago, Chile (e-mail: dsaez@ing.uchile.cl)*
*** *Faculty of Electrical and Computer Engineering, University of Ljubljana, Slovenia, (e-mail: igor.skrjanc@fe.uni-lj.si)*

Abstract: In this paper a new identification method for non-linear hybrid systems that have mixed continuous and discrete states by using fuzzy clustering and principal component analysis is described. The method first determines the hybrid characteristic of the system inspired by an inverse form of the merge method for clusters, which makes it possible to identify the unknown switching points of a process based on just input-output data. Using the switching points, a hard partition of the input-output space is obtained. Then, we propose to use Takagi-Sugeno (TS) fuzzy models with Gaussian MFs as sub-models for each partition. Thus, the overall model is hybrid-fuzzy and will include explicitly the hybrid behavior of the system (the detected switching points) by means of binary MFs, and in each partition all the other non-linearities by means of TS sub-models. An illustrative experiment on a hybrid-tank system is conducted to present the benefits of the proposed approach.

Keywords: Nonlinear system identification; hybrid and distributed system identification; fuzzy identification; fuzzy clustering; principal component analysis.

1. INTRODUCTION

Hybrid systems represent an important class of dynamical systems that contain continuous and discrete/integer variables. Different types of models can be used to represent hybrid systems, for example mixed lineal dynamical models (MLD), linear complementarity, extended linear complementarity, piece-wise affine (PWA), and max-min plus scaling systems. Each sub-class has its own advantages over the others. For example, control techniques for MLD hybrid models, stability criteria for PWA systems, and conditions of existence and uniqueness of solution trajectories for linear complementarity systems, see Bemporad and Morari (1999), Heemels et al. (2001) and references within.

For non-linear systems, there are many identification methodologies such as fuzzy and neural networks modeling. However, few methodologies consider non-linear modeling with continuous and discrete variables, i.e., identification of hybrid systems. The identification methods for hybrid systems are mainly focused on Piecewise ARX (PWARX) systems. See for example Ferrari-Trecate et al. (2003), Juloski et al. (2005i), Bemporad et al. (2005), Gegundez et al. (2008), Drulhe et al. (2008). In Juloski et al. (2005ii) a nice comparison between some of those methods is presented, and in Camacho et al. (2010) a review of identification methods for hybrid systems can be found. In Lauer et al. (2010) a nonlinear hybrid system identification is proposed using kernel functions in order to estimate arbitrary nonlinearities without prior knowledge.

Although most of the developments have been made in conventional fuzzy system a few hybrid fuzzy identification methods have been proposed. Palm and Driankov (1998) presented a hierarchical identification for fuzzy switched systems. The proposed method considers a black-box fuzzy identification by using fuzzy clustering and measurable discrete states in order to obtain a model for continuous state and discrete transitions. Next, Girimonte and Babuška (2004) described two structure-selecting methods for non-linear models with mixed discrete and continuous inputs. The results show that fuzzy clustering is faster in terms of computation time.

In this paper, we propose a new identification method for non-linear hybrid systems that identifies first the discrete transitions (switching points) and then all other kind of non-linearities by only using input-output data of the process, where prior knowledge of discrete modes is not required. The outline of the paper is as follows. In Section 2, the hybrid fuzzy modeling and the identification problem are presented. In Section 3, an identification method

⋆ This research has been supported by the European 7th framework STREP project "Hierarchical and distributed model predictive control (HD-MPC)", contract number INFSO-ICT-223854. The European 7th Framework Network of Excellence "Highly complex and networked control systems (HYCON2)". Also by Fondecyt Chile Grants 1100239-111047, and by the Ministry of Science, Higher Education and Technology of the Republic of Slovenia.

based on fuzzy clustering and the principal components, is presented. Section 4 shows the results of the proposed hybrid fuzzy modeling for a hybrid tank system. Lastly, Section 5 presents the conclusions and further research.

2. PROBLEM STATEMENT

For the modeling of hybrid systems the most popular model types used in the literature are piecewise affine (PWA) system and mixed logical and dynamical (MLD) system. In this paper we propose the use of another type of model called hybrid-fuzzy system, which combines the characteristics of fuzzy models to represent nonlinearities, and the hybrid system to include quantized variables.

We consider hybrid discrete-time nonlinear dynamic systems with input $\mathbf{u}(t) \in R^m$, and to explain the identification method we consider a single output $y(t) \in R$ (the method is easily extensible for multiple outputs). Let $\mathbf{u}^{t-1} = \left[\mathbf{u}(t-1)^T, \ldots, \mathbf{u}(t-n_b)^T\right]^T$, and $\mathbf{y}^{t-1} = [y(t-1), \ldots, y(t-n_a)]^T$ be, respectively, past inputs and outputs up to time $t-1$, n_a and n_b are the model orders. We will assume that the discrete dynamics (transitions) of the system occur when $\mathbf{y}^{t-1}$ satisfies some conditions, and they will not depend on the inputs. This type of hybrid systems is described in general form as:

$$
\begin{aligned}
y(t) &= \sum_{i=1}^{s} f_i(\mathbf{y}^{t-1}, \mathbf{u}^{t-1}) \varrho_i(\mathbf{y}^{t-1}), \\
\varrho_i(\mathbf{y}^{t-1}) &= \begin{cases} 1, & \text{if} \quad \mathbf{y}^{t-1} \in \chi_i \\ 0, & \text{otherwise} \end{cases},
\end{aligned} \tag{1}
$$

where s is the number of discrete modes (sub-models). The local behavior of the system is described by the functions $f_i(\cdot)$ and the discrete mode $\varrho_i(\mathbf{y}^{t-1})$ is a binary variable. The regions χ_i form a complete partition of the output regressor set χ, i.e., $\bigcup_{i=1}^{s} \chi_i = \chi$ and $\chi_i \cap \chi_j = \emptyset$, $\forall i \neq j$.

The aim in this work is to present a systematic method for determining the functions $f_i(\cdot)$ and the regions χ_i given only the input-output data of the process. The functions $f_i(\cdot)$ could be any non-linear function that will be identified by the TS models and the regions χ_i are assumed to be convex polyhedra, described by

$$
\chi_i = \{\mathbf{y}^{t-1} \in R^{n_a} : \quad H_i \mathbf{y}^{t-1} \preceq h_i\} \tag{2}
$$

where $H_i \in R^{q_i \times n_a}$, $h_i \in R^{q_i}$ $i = 1, ..., s$, and $\preceq$ denotes componentwise inequality, where some inequalities are strict to prevent the boundaries of the regions from overlapping. The number of linear inequalities defining the i-th polyhedral region is q_i. In this paper, as a consequence of the algorithm, the resulting H_i, $i = 1, ..., \overline{s}$, are diagonal matrices, so the partition will be a hyperrectangle.

The system given by (1) can be represented by a two-level fuzzy model, which was described by Tanaka et al. (2001). The corresponding two levels are the local fuzzy level and the discrete/quantized level. The local fuzzy level is a set of TS fuzzy models with local validity in one region of an estimated partition $\overline{\chi}_i$, $i = 1, ..., \overline{s}$, where $\overline{s}$ is the estimated number of regions. The discrete/quantized level is given by a set of crisp functions $\delta_i(\mathbf{y}^{t-1})$, which activate the i-th local TS model if $\mathbf{y}^{t-1}$ is in $\overline{\chi}_i$.

Let us assume that input-output data $(y(t), \mathbf{y}^{t-1}, \mathbf{u}^{t-1})$, $t = 1, ..., N$ is available . The structure of a hybrid-fuzzy model to be identified for the variable $y(t)$ is described as:

$$
\begin{aligned}
y(t) &= \sum_{i=1}^{\overline{s}} f_i^{\mathrm{TS}}(\mathbf{z}^{t-1}, \mathbf{y}^{t-1}, \mathbf{u}^{t-1}) \delta_i(\mathbf{y}^{t-1}), \\
\delta_i(\mathbf{y}^{t-1}) &= \begin{cases} 1, & \text{if} \quad \mathbf{y}^{t-1} \in \overline{\chi}_i \\ 0, & \text{otherwise} \end{cases}, \\
f_i^{\mathrm{TS}}(\mathbf{z}^{t-1}, \mathbf{y}^{t-1}, \mathbf{u}^{t-1}) &= \sum_{j=1}^{R_i} \beta_{ij}(\mathbf{z}^{t-1}) y_{ij}(\mathbf{y}^{t-1}, \mathbf{u}^{t-1}), \\
y_{ij}(\mathbf{y}^{t-1}, \mathbf{u}^{t-1}) &= (\mathbf{a}_{ij})^T \mathbf{y}^{t-1} + (\mathbf{b}_{ij})^T \mathbf{u}^{t-1} + r_{ij}, \\
\beta_{ij}(\mathbf{z}^{t-1}) &= \frac{\prod_{r=1}^{p} A_{ij,r}(z_r(t-1))}{\sum_{j=1}^{R_i} \prod_{r=1}^{p} A_{ij,r}(z_r(t-1))},
\end{aligned} \tag{3}
$$

where p is the number of inputs at the premises, the vector of the premises is $\mathbf{z}^{t-1} = [z_1(t-1), \ldots, z_p(t-1)]^T$ and are permitted to be inputs, outputs. We will assume $\mathbf{z}^{t-1} = \left[(\mathbf{y}^{t-1})^T, (\mathbf{u}^{t-1})^T\right]^T$, so $p = n_a + m \cdot n_b$. The index i represents the ith region, $(\mathbf{a}_{ij})^T$, $(\mathbf{b}_{ij})^T$, r_{ij} are the parameters of the fuzzy model $f_i^{\mathrm{TS}}(\cdot)$ for the region i on rule j, R_i is the number of rules of the fuzzy model at the ith region, $A_{ij,r}(z_r(t-1))$ is the membership degree for the input $z_r(t-1)$ at the ith region and rule j, and $\beta_{ij}(\mathbf{z}(t-1))$ is the activation degree of the jth rule that belongs to the fuzzy model of the ith region.

Note also that the model given by (3) is a Takagi-Sugeno fuzzy model, with $\overline{s} \cdot R_i$ rules and activation degree $\beta_{ij}(\mathbf{z}^{t-1})\delta_i(\mathbf{y}^{t-1})$. One of the most important part of the hybrid-fuzzy model is the fuzzy rule base. The rule i, j is the following:

R_{ij} : ***if*** $\boldsymbol{y}^{t-1} \in \overline{\chi}_i$ ***and*** $z_1(t-1)$ *is* $A_{ij,1}$ ***and*** $z_2(t-1)$ *is* $A_{ij,2}$ ***and*** .... ***and*** $z_p(t-1)$ *is* $A_{ij,p}$ ***then*** $y_{ij}(t) = (\boldsymbol{a}_{ij})^T \boldsymbol{y}^{t-1} + (\boldsymbol{b}_{ij})^T \boldsymbol{u}^{t-1} + r_{ij}, \quad j = 1, .., R_i, \quad i = 1, .., \overline{s}$.

Note that the first component $\mathbf{y}^{t-1} \in \overline{\chi}_i$ of the fuzzy rule evaluates the binary membership function $\delta_i(\mathbf{y}^{t-1})$ and it explicitly incorporates the discrete transitions of the system.

By only using a finite input-output data set of the process, the identification problem of a hybrid-fuzzy model given by (3) consists of estimating the following parameters: $\overline{s}$, $\overline{\chi}_i$, $i = 1, .., \overline{s}$, R_i, $A_{ij,r}(\cdot)$, $(\mathbf{a}_{ij})^T$,$(\mathbf{b}_{ij})^T$, and r_{ij}. As explained in Bemporad et al. (2005), usually an identification procedure is carried out by minimizing a cost function:

$$
V_N = \frac{1}{N} \sum_{t=1}^{N} J \left(y(t) - \sum_{i=1}^{\overline{s}} f_i^{\mathrm{TS}}(\mathbf{z}^{t-1}, \mathbf{y}^{t-1}, \mathbf{u}^{t-1}) \delta_i(\mathbf{y}^{t-1}) \right) \tag{4}
$$

where J is a penalty function for the error, typically a quadratic function. The optimization problem should also include additional terms if we want to avoid overfit, or if we want meaningful fuzzy rules. The minimization of (4) is in general a non-convex non-linear mixed-integer optimization problem. In this paper we propose a new method for the identification of hybrid-fuzzy systems based on well-known principles, that identifies first the discrete transitions and then all other kind of non-linearities.

3. HYBRID-FUZZY IDENTIFICATION METHOD

A new fuzzy-hybrid identification method is developed based on fuzzy clustering and principal component. The method first allows to identify the discrete transitions (switching points) and then all other kind of non-linearities by only using input-output data, where prior knowledge of the discrete modes is not required.

As a motivation example, let us consider the hybrid tank system in Figure 1. In the figure, A_1, A_2 and A_3 are the cross-section of the tanks, S_1, S_2 and S_3 are the cross-section of the outlet holes, g is the acceleration due to gravity, $Q(t)$ is the input flow, and $h(t)$ is the level of the tank. The hybrid tank system is divided into three region because the cross-section of the tank is larger when the level is higher than h_1 and h_2.

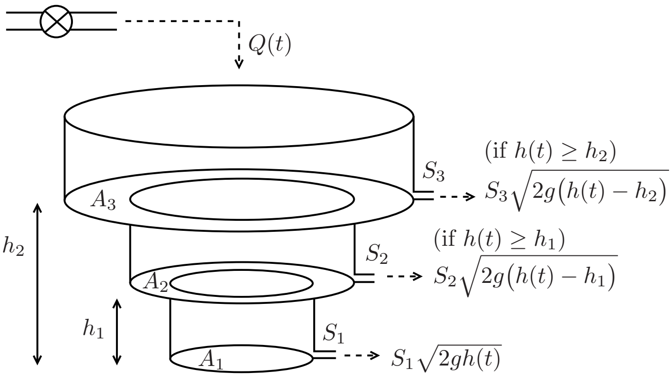

Fig. 1. Hybrid Tank System

Thus, for a fixed input flow, it will take more time to increase the level when it is higher than h_1 or h_2, as compared $h(t)$ when it is lower, because the cross-section is larger. This means that the level values h_1 and h_2 are switching points in the sense that those levels are the border of the three different operating regions, the dynamics of which are different. Next, the hybrid-fuzzy identification method is presented.

3.1 Hybrid-fuzzy Model Identification Procedure

Throughout this paper we assume that N input/output data have been collected:

$$\Phi = \begin{bmatrix} y(1) & (\mathbf{y}^0)^T & (\mathbf{u}^0)^T \\ y(2) & (\mathbf{y}^1)^T & (\mathbf{u}^1)^T \\ \vdots & \vdots & \vdots \\ y(N) & (\mathbf{y}^{N-1})^T & (\mathbf{u}^{N-1})^T \end{bmatrix}_{N, n_a+m\cdot n_b+1}, \quad (5)$$

where N denote the number of data samples, $y(t) \in R$ is the variable we want to estimate with the hybrid-fuzzy model, $\mathbf{y}^{t-1} \in R^{n_a}$ are past outputs up to time $t-1$, $\mathbf{u}^{t-1} \in R^{m\cdot n_b}$ are past inputs up to time $t-1$, and n_a and n_b are the model orders.

The identification procedure consists of the following seven steps:

Step 1: Determine the fuzzy clusters over the data Φ, using the G-K algorithm Gustafson and Kessel (1978). This algorithm searches for hyperplanes in an n-dimensional space. Then, it is suitable for the identification of hybrid-fuzzy models because the consequents of hybrid-fuzzy models are hyperplanes in the premise-consequent product space. The algorithm will cluster the data given a specified number of cluster c and the parameters for the cluster fuzziness and the stopping criterion. The G-K algorithm provides the centers of clusters $\mathbf{v}_l = [v_l^1, ..., v_l^{n_a+m\cdot n_b+1}]^T$, a covariance matrix for each fuzzy cluster l, with $n_a + m \cdot n_b + 1$ eigenvectors $\{\varphi_{1,l}, \ldots, \varphi_{n_a+m\cdot n_b+1,l}\}$, and with the corresponding eigenvalues $\{\lambda_{1,l}, \ldots, \lambda_{n_a+m\cdot n_b+1,l}\}$.

It is well known that G-K algorithm does not give an indication of the correct number of clusters c needed. A large number of clusters will result in a complicated rule-base model, while a small number of clusters result in a poor model. It is also important to preserve the small clusters in the interesting regions, which may have been found when clustering with an initially large number of clusters. So to obtain the optimum number of clusters we propose using the compatible cluster merging method, just like it is suggested for the identification of TS models in Babuška (1998), Kaymak and Babuška (1995).

Step 2: To select the eigenvector $\varphi_l^* = [\phi_l^1, ..., \phi_{l_1}^{n_a+m\cdot n_b+1}]^T$ associated with the maximum eigenvalue λ_l^* for each cluster $l = 1, \ldots, c$:

$$\lambda_l^* = \max\{\lambda_{1,l}, \lambda_{2,l}, \ldots, \lambda_{n_a+m\cdot n_b+1,l}\}. \quad (6)$$

We propose to detect the switching points by analyzing the most important eigenvectors (the principal vectors or the principal components), in which directions the most information is given. Inspired by the merge method for clusters Kaymak and Babuška (1995), we will look for clusters whose centers are sufficiently close (consecutive clusters), but instead of merging parallel hyperplanes clusters, we will split the output-regressor space when those consecutive clusters are very different (angle between the hyperplanes is big). We assume that the switching points are in the outputs, so the analysis will be done for each component of the output-regressor space $y(t-k)$, $k = 1, ..., n_a$.

Step 3: For every cluster $l = 1, ..., c$ and every component of the output regressor space $y(t-k)$, $k = 1, ..., n_a$, to calculate the vector $\hat{\pi}_{lk}$, which represents the projection of the eigenvector φ_l^* on the subspace given by the inputs and the output $y(t-k)$, and which is given by:

$$\hat{\pi}_{lk} = \frac{\Phi_k \varphi_l^*}{\|\Phi_k \varphi_l^*\|_2}, \quad \forall l \in \{1, ..., c\}, \quad \forall k \in \{1, ..., n_a\}, \quad (7)$$

where φ_l^* is the eigenvector chosen in step 2 and Φ_k is the matrix dimension $(n_a + m \cdot n_b + 1) \times (n_a + m \cdot n_b + 1)$, the elements of which are defined as:

$$(\Phi_k)_{\ell,\wp} = \begin{cases} 1 & \text{if} \quad \ell = \wp = k+1, \\ 1 & \text{if} \quad \ell = \wp \quad \text{and} \quad \ell > n_a + 1, \\ 0 & \text{if} \quad \text{otherwise.} \end{cases} \tag{8}$$

Note that the vector is normalized, so $\|\hat{\pi}_{lk}\|_2 = 1$.

Step 4: For every vector $\hat{\pi}_{lk}$ determine $\hat{\pi}_{lk}^u$ which represents the projection of $\hat{\pi}_{lk}$ in the subspace generated by the inputs, and which is obtained in the following way:

$$\hat{\pi}_{lk}^u = \frac{\Phi_u \hat{\pi}_{lk}}{\|\Phi_u \hat{\pi}_{lk}\|}, \quad \forall l \in \{1, ..., c\}, \quad \forall k \in \{1, ..., n_a\}, \tag{9}$$

where $\hat{\pi}_{il}$ is the vector obtained in *Step 3*, and Φ_u is the matrix of dimension $(n_a + m \cdot n_b + 1) \times (n_a + m \cdot n_b + 1)$, the elements of which are defined as:

$$(\Phi_u)_{\ell,\wp} = \begin{cases} 1 & \text{if} \quad \ell = \wp \quad \text{and} \quad \ell > n_a + 1, \\ 0 & \text{if} \quad \text{otherwise.} \end{cases} \tag{10}$$

Note that the vector is normalized, so $\|\hat{\pi}_{lk}^u\| = 1$. Finally, for each cluster l and every output variable $y(t-k)$, compute the cluster slope $\Gamma_{lk} = \tan(\hat{\gamma}_{lk})$ given by:

$$\Gamma_{lk} = \sqrt{\frac{1}{(\hat{\pi}_{lk}^T \hat{\pi}_{lk}^u)^2} - 1}, \quad \forall l \in \{1, ..., c\}, \quad \forall k \in \{1, ..., n_a\}, \tag{11}$$

Step 5: In this step the idea is to determine possible switching points for every variable $y(t-k)$. For doing this, if l_1 and l_2 are two consecutive clusters, with center $\mathbf{v}_{l_1}$ and $\mathbf{v}_{l_2}$ in descending order for the variable $y(t-k)$, $(v_{l_1}^{k+1} < v_{l_2}^{k+1})$, to evaluate the rate $\Delta\Gamma_{l_1 l_2 k}$ given by:

$$\Delta\Gamma_{l_1 l_2 k} = |\Gamma_{l_1 k} - \Gamma_{l_2 k}|. \tag{12}$$

The candidate switching point should be in between the coordinates $v_{l_1}^{k+1}$ and $v_{l_2}^{k+1}$. We propose estimating the location of the switching point $V_k^{l_1 l_2}$ in the following way:

$$V_k^{l_1 l_2} = \frac{\frac{v_{l_1}^{k+1} + \sqrt{\lambda_{l_1}^*}\phi_{l_1}^{k+1}}{\lambda_{l_1}^*} + \frac{v_{l_2}^{k+1} - \sqrt{\lambda_{l_2}^*}\phi_{l_2}^{k+1}}{\lambda_{l_2}^*}}{\frac{1}{\lambda_{l_1}^*} + \frac{1}{\lambda_{l_2}^*}}. \tag{13}$$

where $\lambda_{l_1}^*$ and $\lambda_{l_2}^*$ are the eigenvalues obtained in *Step 2* corresponding to clusters l_1 and l_2 respectively and $\phi_{l_1}^{k+1}$ and $\phi_{l_2}^{k+1}$ are the $k+1$-th coordinates of the corresponding eigenvectors.

The next step is to choose the switching point candidates $V_k^{l_1 l_2}$ the rate of which $\Delta\Gamma_{l_1 l_2 k}$ satisfies a criterion. A sensitivity analysis could be performed to evaluate whether the inclusion of a switching point improves the performance of the prediction model or not. Then, we will add one switching point, we will identify the hybrid-fuzzy model, and then we will analyze again *Step 5*, to determine the inclusion of another switching point. The process will finish once the performance of the hybrid-fuzzy model is not improved significantly by the inclusion of new switching points. So, let assume we have generate a partition $\{\overline{\chi}_i\}_{i=1}^{\overline{s}}$. We now analyze the inclusion of a new switching point in the model, by splitting the region $\overline{\chi}_i$ into two new regions divided by the new switching point. So, let consider the switching point candidate, with the maximum rate, given by:

$$V_{\overline{s}} = \{V_{\overline{k}}^{\overline{l_1 l_2}} : \quad (\overline{l_1}, \overline{l_2}, \overline{k}) = \text{argmax}\{\Delta\Gamma_{l_1 l_2 k}\}\}. \tag{14}$$

Step 6: To split the region $\overline{\chi}_i$ into two new regions. Recall that the region $\overline{\chi}_i$ is defined as follows:

$$\overline{\chi}_i = \left\{\mathbf{y}^{t-1} : \quad H_i \mathbf{y}^{t-1} \preceq h_i\right\},$$

where $H_i \in R^{q_i \times n_a}$, $h_i \in R^{q_i}$ $i = 1, ..., \overline{s}$, the symbol $\preceq$ denotes componentwise inequality, where some inequalities are strict to avoid the boundaries of the regions to have multiple values. Given the new switching point $V_{\overline{s}}$ in the variable $y(t-k)$, the two new regions are defined as follows:

$$\overline{\chi}_1^i = \left\{\mathbf{y}^{t-1} : \quad H_i \mathbf{y}^{t-1} \preceq h_i \quad \wedge \quad y(t-k) \leq V_{\overline{s}}\right\}.$$

$$\overline{\chi}_2^i = \left\{\mathbf{y}^{t-1} : \quad H_i \mathbf{y}^{t-1} \preceq h_i \quad \wedge \quad -y(t-k) < -V_{\overline{s}}\right\}.$$

Step 7: For the sub-regions $\overline{\chi}_1^i$ and $\overline{\chi}_2^i$, a local TS model is identified. First, we split the data belonging to the region $\overline{\chi}_i$ into the two new regions, by the rule: if $y(t-k) \leq V_{\overline{s}}$ then $(y(t), (\mathbf{y}^{t-1})^T, (\mathbf{u}^{t-1})^T) \in \overline{\chi}_1^i$, else $(y(t), (\mathbf{y}^{t-1})^T, (\mathbf{u}^{t-1})^T) \in \overline{\chi}_2^i$, $t = 1, ..., N$. Then, for each new partition, just considering data that belongs to the sub-region, the number of rules R_i and the membership functions $A_{ij,r}(\cdot)$ are obtained with a clustering method (G-K). Each TS model is optimized for the number of fuzzy clusters and their regressor structure is obtained by a sensitivity analysis.

The next step is to identify the consequent parameters of each rule of the TS model, see Karer et al. (2007). Let us write the consequent parameters for the fuzzy rule j in the region i as follow:

$$\Theta_{ij} = \begin{bmatrix} \mathbf{a}_{ij} \\ \mathbf{b}_{ij} \\ r_{ij} \end{bmatrix}_{n_a + m \cdot n_b + 1, 1}, \tag{15}$$

The model parameters for the rule j of region i can be obtained using the least-squares identification method as follows:

$$\Theta_{ij} = (\Psi_{ij}^T \Psi_{ij})^{-1} \Psi_{ij}^T Y_{ij} \tag{16}$$

where the matrices Ψ_{ij} and Y_{ij} are the following:

$$\Psi_{ij} = \begin{bmatrix} \beta_{ij}(\mathbf{z}^0)[(\mathbf{y}^0)^T \quad (\mathbf{u}^0)^T \quad 1] \\ \beta_{ij}(\mathbf{z}^1)[(\mathbf{y}^1)^T \quad (\mathbf{u}^1)^T \quad 1] \\ \vdots \\ \beta_{ij}(\mathbf{z}^{N_{ij}-1})[(\mathbf{y}^{N_{ij}-1})^T \quad (\mathbf{u}^{N_{ij}-1})^T \quad 1] \end{bmatrix}, \tag{17}$$

$$Y_{ij} = \begin{bmatrix} \beta_{ij}(\mathbf{z}^0)y(1) \\ \beta_{ij}(\mathbf{z}^1)y(2) \\ \vdots \\ \beta_{ij}(\mathbf{z}^{N_{ij}-1})y(N_{ij}) \end{bmatrix}_{N_{ij},1}, \tag{18}$$

where N_{ij} is the number of input-output data pairs corresponding to the rule j of the region i considering only the data that belongs to the region i and $\beta_{ij}(\mathbf{z}(t-1)) \geq \delta$, with δ a small positive number essential for obtaining suitable conditioned matrices.

4. SIMULATION RESULTS

4.1 Hybrid Tank System

Let us consider the hybrid tank system shown in Figure 1, a modification of the one used in Gegundez et al. (2008). The following non-linear equations describe the dynamic of the tank system:

$$\frac{dh}{dt} = \begin{cases} \dfrac{1}{A_1}(Q(t) - F_1(t)) & \text{if} \quad h(t) \leq h_1 \\ \dfrac{1}{A_2}(Q(t) - F_1(t) - F_2(t)) & \text{if} \quad h_1 < h(t) \leq h_2 \\ \dfrac{1}{A_3}(Q(t) - F_1(t) - F_2(t) - F_3(t)) & \text{if} \quad h(t) \geq h_2 \end{cases}, \tag{19}$$

where $h(t)[m]$ is the level of the tank, $u(t) = Q(t)[m^3/s]$ is the input flow, outflows are $F_1(t) = S_1\sqrt{2gh(t)}$, $F_2(t) = S_2\sqrt{2g(h(t) - h_1)}$, $F_3(t) = S_3\sqrt{2g(h(t) - h_2)}$, $A_1 = 0.0154[m^2]$ is the cross-section of the first region of the tank, the cross-section of the second and third regions are given by $A_2 = 3A_1$, $A_3 = 9A_1$, $S_1 = S_2 = S_3 = 0.0005[m^2]$ are the cross-section of the outlet holes, and $g = 9.81[m^2/s]$ is the acceleration due to gravity. The hybrid tank system is divided into three regions because the cross-section of the tank is three times bigger when the level is higher than $h_1 = 0.2[m]$ and then three time bigger when the level is higher than $h_2 = 0.4[m]$. The identification problem is to find the relation between $h(t)$ and $Q(t)$ considering the input/output data. The main goal is to find the number of switching regions and the switching point (in this case $h(t) = 0.2$ and $h(t) = 0.4$), which defines the partition. The input/output data considered are $\mathbf{y}^{t-1} = h(t-1)$ as the output and $\mathbf{u}^{t-1} = Q(t-1)$ as the input. In order to evaluate the performance of the hybrid-fuzzy model (with one and two switching points detected) and TS model (with no switching point included), the Root Mean Squared (RMS) error is used.

The signals were sampled with $T_s = 10[s]$. For the input a uniformly distributed random signal with minimum value 0 and maximum value 0.005 was used. A total of 1000 samples were used as training set, and 1000 as the validation set.

For the hybrid-fuzzy models, a switching point was estimated to be in $h(t-1) = 0.385[m]$ (the real value is 0.4). After splitting the data into the new regions $h(t-1) \geq 0.385$ and $h(t-1) < 0.385$, the rates between consecutive clusters belonging to each region are calculated again, the switching point being estimated to be in $h(t-1) = 0.25[m]$ (the real value is 0.2). There are three subregions ($\overline{s} = 3$): The first one is $\overline{\chi}_{21}$, where $h(t-1) < 0.25$. The second is $\overline{\chi}_{22}$, where $h(t-1) < 0.385$ and $h(t-1) \geq 0.25$. The third is $\overline{\chi}_{23}$, where $h(t-1) \geq 0.385$. The structure of hybrid-fuzzy model is given by:

R_{ij} : ***if*** $h(t-1) \in \overline{\chi}_i$ ***and*** $h(t-1)$ *is* $A_{ij,1}$ ***and*** $Q(t-1)$ *is* $A_{ij,2}$, ***then*** $h_{ij}(t) = a_{ij}h(t-1) + b_{ij}Q(t-1) + r_{ij}$, $j = 1, .., R_i$,

where $A_{ij,r}(z_r(t-1)) = e^{-0.5(c_{1,ij,r}(z_r(t-1)-c_{2,ij,r}))^2}$. The parameters for hybrid-fuzzy model is given in Table 1.

Table 1. Parameters of hybrid-fuzzy model-2

$\overline{\chi}_{23}$

j	$c_{1,1j,1}$	$c_{2,1j,1}$	$c_{1,1j,2}$	$c_{2,1j,2}$	a_{1j1}	b_{1j1}	r_{1j}
1	4.484	183.27	0.4689	0.0023	-0.1617	0.771	81.02
2	5.360	153.32	0.4403	0.0022	-0.1675	0.777	81.95
3	5.118	160.56	0.5224	0.0026	-0.1572	0.767	79.68

$\overline{\chi}_{22}$

j	$c_{1,2j,1}$	$c_{2,2j,1}$	$c_{1,2j,2}$	$c_{2,2j,2}$	a_{2j1}	b_{2j1}	r_{2j}
1	7.823	129.14	0.3513	0.0018	0.0213	0.211	98.24
2	4.712	214.38	0.2462	0.0012	0.0119	0.224	101.43
3	4.440	227.54	0.2857	0.0014	0.0137	0.221	101.07

$\overline{\chi}_{21}$

j	$c_{1,3j,1}$	$c_{2,3j,1}$	$c_{1,3j,2}$	$c_{2,3j,2}$	a_{3j1}	b_{3j1}	r_{3j}
1	6.679	146.86	0.0933	0.0005	0.0539	-0.036	102.64
2	5.555	176.58	0.0277	0.0001	0.0531	-0.043	103.58
3	5.623	174.44	0.1562	0.0008	0.0511	-0.033	103.95
4	9.106	107.71	0.1488	0.0007	0.0515	-0.025	102.77

Comparative Analysis. Table 2 contains the RMS errors divided by the number of data points for the hybrid-fuzzy (H-F) and TS models, considering the validation data set for 1, 5, and 10, step-ahead prediction. Figure 2 show the measured output and the output predicted by H-F model with two switches (H-F model 2).

Table 2. RMS error, TS and hybrid-fuzzy model, validation data.

STEPS.	TS	H-F1	H-F2
1	0.0405	0.0354	0.0353
5	0.0420	0.0378	0.0363
10	0.0448	0.0395	0.0378

A switching point was detected at $h(t-1) = 0.385$. In the case of $h(t-1)$, the real switching point was set to 0.4 $[m]$, which is a pretty fairly good estimation. The detection of the second switch was more difficult, because of the effect of the switching point at $h(t-1) = 0.385$ in the clusters close to the border. From the Figure 2, and Table 2, we can say that the main advantage of hybrid-fuzzy modeling is its fuzzy rules, which can be used directly to detect the modes of the system.

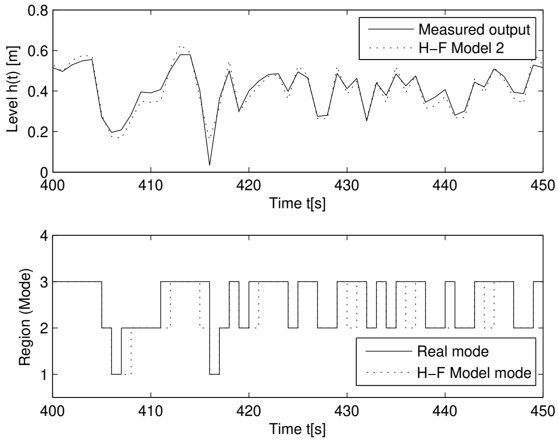

Fig. 2. Measured and predicted (a) output, (b) and modes using H-F Model 2

5. CONCLUSION

In this paper a new identification method for non-linear hybrid systems that identifies discrete transitions by using only input-output data has been presented. A hybrid-fuzzy model was identified, which consist of a local fuzzy level and a discrete/quantized level. Thus, the hybrid-fuzzy model incorporates explicitly the hybrid behavior of the process. Moreover, the method was implemented and applied to a tank-system. The algorithm is a mixture of existing methods (principal component analysis, fuzzy clustering) and demonstrated to be very useful in the detection of switching points by simulation. The comparisons demonstrated the better performance of hybrid-fuzzy models compared to the conventional TS model when comparing prediction performance. However, we must point out that the main advantage of hybrid-fuzzy modeling are the rules with explicit information about the modes of the plant.

In further research, new approaches of hybrid-fuzzy modeling will be analyzed such as a fuzzy clustering that generates both the fuzzy and hard partitions. The stability issues of the proposed hybrid-fuzzy models can also be studied. State-space model identification and estimation is also an interesting topic for this class of non-linear systems. Online clustering, or learning methods could be also applied in a further stage.

ACKNOWLEDGEMENTS

We thanks Gorazd Karer and Patricio Torres for their contributions in the initial stages of this research.

REFERENCES

R. Babuška, *Fuzzy Modelling for Control*, KAP, 1998.

A. Bemporad and M. Morari, *Control of systems integrating logic, dynamics and constraints*, Automatica, vol. 35(3), pp. 407-427, 1999.

A. Bemporad, A. Garulli, S. Paoletti and A. Vicino, *A bounded-error approach to piecewise affine system identification*, IEEE Trans on Automatic Control, vol. 50, pp. 1567-1580, 2005.

E.F. Camacho, D.R. Ramirez, D. Limon, D. Munoz de la Pena and T. Alamo, *Model Predictive control techniques for hybrid systems*, Annual Reviews in Control, vol. 34, pp. 21-31, 2010.

S. Drulhe, G. Ferrari-Trecate and H. de-Jong, *The switching threshold reconstruction problem for piecewise-affine models of genetic regulatory networks*, IEEE Transactions on Automatic Control, vol. 53, pp. 153-165, 2008.

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

M.E. Gegundez, J. Aroba and J.M. Bravo, *Identification of piecewise affine systems by means of fuzzy clustering and competitive learning*, Engineering Applications of Artificial Intelligence, vol. 21, pp. 1321-1329, 2008.

D. Girimonte and R. Babuška, *Structure for nonlinear models with mixed discrete and continuous inputs: a comparative study*, In: Proc of IEEE International Conference on Systems, Man and Cybernetics, pp. 2392-2397, 2004.

D.E. Gustafson and W.C. Kessel, *Fuzzy clustering with a fuzzy covariance matrix*, In: Proc IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes, pages 761-766, San Diego, CA, USA, 1978.

W.P.M.H. Heemels, B. De Schutter and A. Bemporad, *Equivalence of hybrid dynamical models*, Automatica, Vol. 37, pp. 1085-1091, 2001.

A. Juloski, S. Weiland and W. Heemels, *A Bayesian approach to identification of Hybrid Systems*, IEEE Transactions on Automatic Control, vol. 50, pp. 1520-1533, 2005i.

A. Juloski, W.P.M.H. Heemels, G. Ferrari-Trecate, R. Vidal, S. Paoletti and J.H.G. Niessen. *Comparison of four procedures for the identification of hybrid systems*, In M. Morari and L. Thiele, editors, Hybrid Systems: Computation and Control vol. 3414 of Lecture Notes in Computer Science, pp. 354-369. Springer Verlag, 2005ii.

G. Karer, G. Music, I. Škrjanc and B. Zupančič, *Hybrid fuzzy model-based predictive control of temperature in a batch reactor*, Computers & Chemical Engineering, vol. 31, pp. 1552-1564, 2007.

U. Kaymak and R. Babuška, *Compatible cluster merging for fuzzy modeling*, Proceedings of FUZZ-IEEE/IFES95, Yokohama, Japan, pp. 897-904, 1995.

F. Lauer, G. Bloch and R. Vidal, *Nonlinear hybrid system identification with kernel models*, Proceedings of 49th IEEE Conference on Decision and Control (CDC), pp. 696-701, 2010. 897-904, 1995.

R. Palm and D. Driankov, *Fuzzy switched hybrid systems - Modeling and identification*, In: Proc of the 1998 IEEE ISCI/CIRA/SAS Joint Conf, Gaithersburg MD, USA, pp. 130-135, Sept. 1998.

T. Takagi and M. Sugeno, *Fuzzy identification of systems and its applications to modeling and control*, IEEE Transactions on Systems, Man, Cybernetics, vol. 15, pp. 116-132, 1985.

K. Tanaka, M. Iwasaki and H. Wang, *Switching control of an R/C hovercraft: stabilization and smooth switching*, IEEE Transactions on Systems, Man, Cybernetics, vol. 31, pp. 853-863, 2001.