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 variables. Different types of models can be used to represent hybrid systems, for example mixed linear 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
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 a different 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 \( u(t) \in \mathbb{R}^m \), and to explain the identification method we consider a single output \( y(t) \in \mathbb{R} \) (the method is easily extendible for multiple outputs). Let \( u^{t-1} = [u(t-1)^T, \ldots, u(t-n_u)^T]^T \), and \( y^{t-1} = [y(t-1), \ldots, y(t-n_y)]^T \) be, respectively, past inputs and outputs up to time \( t-1 \), \( n_u \) and \( n_y \) are the model orders. We will assume that the discrete dynamics (transitions) of the system occur when \( y^{t-1} \) satisfies some conditions, and they will not depend on the inputs. This type of hybrid systems is described in general form as:

\[
y(t) = \sum_{i=1}^{s} f_i(y^{t-1}, u^{t-1})g_i(y^{t-1}),
\]

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 \( g_i(y^{t-1}) \) is a binary variable. The regions \( \chi_i \) form a complete partition of the output regressor set \( \chi \), 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 \( \chi_i \), 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 \( \chi_i \) are assumed to be convex polyhedra, described by

\[
\chi_i = \{y^{t-1} \in \mathbb{R}^{n_y} : H_i y^{t-1} \preceq h_i \}
\]

where \( H_i \in \mathbb{R}^{q_i \times n_y} \), \( h_i \in \mathbb{R}^{q_i} \), \( i = 1, \ldots, s \), 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 \( H_i, i = 1, \ldots, \Sigma \), 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 \( \chi_i, i = 1, \ldots, \Sigma \), where \( \Sigma \) is the estimated number of regions. The discrete/quantized level is given by a set of crisp functions \( \delta_i(y^{t-1}) \), which activate the \( i \)-th local TS model if \( y^{t-1} \) is in \( \chi_i \).

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

\[
y(t) = \sum_{i=1}^{\Sigma} f_i^{TS}(z^{t-1}, y^{t-1}, u^{t-1})\delta_i(y^{t-1}),
\]

\[
\delta_i(y^{t-1}) = \begin{cases} 1, & \text{if } y^{t-1} \in \chi_i, \\ 0, & \text{otherwise} \end{cases}
\]

\[
f_i^{TS}(z^{t-1}, y^{t-1}, u^{t-1}) = \sum_{j=1}^{R_i} \beta_{ij}(z^{t-1})y_{ij}(y^{t-1}, u^{t-1}),
\]

\[
y_{ij}(y^{t-1}, u^{t-1}) = (a_{ij})^T y^{t-1} + (b_{ij})^T u^{t-1} + r_{ij},
\]

\[
\beta_{ij}(z^{t-1}) = \frac{\prod_{j=1}^{p} A_{ij,r}(z_r(t-1))}{\sum_{j=1}^{R_i} \prod_{j=1}^{p} A_{ij,r}(z_r(t-1))},
\]

where \( p \) is the number of inputs at the premises, the vector of the premises is \( z^{t-1} = [z_1(t-1), \ldots, z_p(t-1)]^T \), and are permitted to be inputs, outputs. We will assume \( z^{t-1} = [(y^{t-1})^T, (u^{t-1})^T]^T \), so \( p = n_u + n_y \). The index \( i \) represents the \( i \)-th region, \( (a_{ij})^T, (b_{ij})^T, r_{ij} \) are the parameters of the fuzzy model \( f_i^{TS} \) for the region \( i \) on rule \( j, R_i \) is the number of rules of the fuzzy model at the \( i \)-th region. \( A_{ij,r}(z_r(t-1)) \) is the membership degree for the input \( z_r(t-1) \) at the \( i \)-th region and rule \( j \), and \( \beta_{ij}(z(t-1)) \) is the activation degree of the \( j \)-th rule that belongs to the fuzzy model of the \( i \)-th region.

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

\[
R_{ij} : \text{if } y^{t-1} \in \chi_i \text{ and } z_j(t-1) = A_{ij,1} \text{ and } z_2(t-1) = A_{ij,2} \text{ and } \ldots \text{ and } z_p(t-1) = A_{ij,p} \text{ then } y_{ij}(t) = (a_{ij})^T y^{t-1} + (b_{ij})^T u^{t-1} + r_{ij}, \quad j = 1, \ldots, R_i, \quad i = 1, \ldots, \Sigma.
\]

Note that the first component \( y^{t-1} \) of \( \chi_i \) of the fuzzy rule evaluates the binary membership function \( \delta_i(y^{t-1}) \) 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: \( \Theta, \Sigma, \chi_i, i = 1, \ldots, \Sigma, R_i, A_{ij,r}(\cdot), (a_{ij})^T, (b_{ij})^T, \) and \( r_{ij} \). As explained in Bemporad et al. (2005), usually an identification procedure is carried out by minimizing a cost function:

\[
V_N = \frac{1}{N} \sum_{t=1}^{N} J \left( y(t) - \sum_{i=1}^{\Sigma} f_i^{TS}(z^{t-1}, y^{t-1}, u^{t-1})\delta_i(y^{t-1}) \right),
\]

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:

\[
\Phi = \begin{bmatrix}
    y(1)^T \\
    (y(2)^T) \\
    \vdots \\
    (y(N)^T) \\
\end{bmatrix}
\begin{bmatrix}
    (u(0)^T) \\
    (u(1)^T) \\
    \vdots \\
    (u(N-1)^T) \\
\end{bmatrix}^{N,n_a+m+n_b+1},
\]

(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, \( y^{t-1} \in R^{n_y} \) are past outputs up to time \( t-1 \), \( u^{t-1} \in R^{m+n_y} \) 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 \( \Phi \), 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 consequences 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 \( \mathbf{v}_l = [v_1^l, \ldots, v_{n_a+m+n_b+1}^l]^T \), a covariance matrix for each fuzzy cluster \( l \), with \( n_a+m+n_b+1 \) eigenvectors \( \{\varphi_{1,l}, \ldots, \varphi_{n_a+m+n_b+1,l}\} \), and with the corresponding eigenvalues \( \{\lambda_1,l, \ldots, \lambda_{n_a+m+n_b+1,l}\} \).

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 \( \varphi^*_l = [\varphi^*_1, \ldots, \varphi^*_{n_a+m+n_b+1}]^T \) associated with the maximum eigenvalue \( \lambda^*_l \) for each cluster \( l = 1, \ldots, c \):

\[
\lambda^*_l = \max\{\lambda_{1,l}, \lambda_{2,l}, \ldots, \lambda_{n_a+m+n_b+1,l}\}. \quad (6)
\]

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-repressor space \( y(t-k), k = 1, \ldots, n_a \).

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

\[
\hat{\pi}_{lk} = \frac{\Phi_k \varphi^*_l}{\|\Phi_k \varphi^*_l\|_2}, \quad \forall l \in \{1, \ldots, c\}, \quad \forall k \in \{1, \ldots, n_a\}, \quad (7)
\]
where \( \varphi^*_l \) is the eigenvector chosen in step 2 and \( \Phi_k \) 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:

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

Note that the vector is normalized, so \( \|\tilde{\pi}_{lk}\|_2 = 1 \).

**Step 4:** For every vector \( \tilde{\pi}_{lk} \) determine \( \tilde{\pi}^n_{lk} \) which represents the projection of \( \tilde{\pi}_{lk} \) in the subspace generated by the inputs, and which is obtained in the following way:

\[
\tilde{\pi}^n_{lk} = \frac{\Phi_k \tilde{\pi}_{lk}}{\|\Phi_k \tilde{\pi}_{lk}\|}, \quad \forall l \in \{1, \ldots, c\}, \quad \forall k \in \{1, \ldots, n_a\}, \tag{9}
\]

where \( \tilde{\pi}_d \) is the vector obtained in Step 3, and \( \Phi_a \) 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:

\[
(\Phi_a)_{\ell, \varphi} = \begin{cases} 1 & \text{if } \ell = \varphi \text{ and } \ell > n_a + 1, \\ 0 & \text{otherwise}. \end{cases} \tag{10}
\]

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

\[
\Gamma_{lk} = \sqrt{\frac{1}{\|\tilde{\pi}^n_{lk}\|^2} - 1}, \quad \forall l \in \{1, \ldots, c\}, \quad \forall k \in \{1, \ldots, n_a\}, \tag{11}
\]

**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 \( v_{l_1} \) and \( v_{l_2} \) in descending order for the variable \( y(t - k) \), \((v^{k+1}_{l_1} < v^{k+1}_{l_2})\), to evaluate the rate \( \Delta \Gamma_{l_1l_2k} \) given by:

\[
\Delta \Gamma_{l_1l_2k} = |\Gamma_{l_1k} - \Gamma_{l_2k}|. \tag{12}
\]

The candidate switching point should be in between the coordinates \( v^{k+1}_{l_1} \) and \( v^{k+1}_{l_2} \). We propose estimating the location of the switching point \( V_{l_1l_2}^{k+1} \) in the following way:

\[
V_{l_1l_2}^{k+1} = \frac{\lambda^{k+1}_{l_1}v^{k+1}_{l_1} + \lambda^{k+1}_{l_2}v^{k+1}_{l_2}}{\lambda^{k+1}_{l_1} + \lambda^{k+1}_{l_2}}. \tag{13}
\]

where \( \lambda^{k}_{l_1} \) and \( \lambda^{k}_{l_2} \) are the eigenvalues obtained in Step 2 corresponding to clusters \( l_1 \) and \( l_2 \) respectively and \( v^{k+1}_{l_1} \) and \( v^{k+1}_{l_2} \) are the \( k \)-th coordinates of the corresponding eigenvectors.

The next step is to choose the switching point candidates \( V_{l_1l_2}^{k} \) the rate of which \( \Delta \Gamma_{l_1l_2k} \) 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 \( \{\tilde{\chi}_i\}_{i=1}^{s} \) and then we will analyze the inclusion of a new switching point in the model, by splitting the region \( \tilde{\chi}_i \) into two new regions divided by the new switching point. So, let consider the switching point candidate, with the maximum rate, given by:

\[
V_{\tilde{\chi}} = \{V_{l_1l_2}^{k+1} : (l_1, l_2, k) = \arg\max\{\Delta \Gamma_{l_1l_2k}\}\}. \tag{14}
\]

**Step 6:** To split the region \( \tilde{\chi}_i \) into two new regions. Recall that the region \( \tilde{\chi}_i \) is defined as follows:

\[
\tilde{\chi}_i = \{y^{t-1} : H_i y^{t-1} \leq h_i\},
\]

where \( H_i \in R^{q \times n_r} \), \( h_i \in R^q \), \( i = 1, \ldots, s \), the symbol “\( \leq \)” denotes componentwise inequality, where some inequalities are strict to avoid the boundaries of the regions to have multiple values. Given the new switching point \( V_{\tilde{\chi}} \) in the variable \( \eta(t - k) \), the two new regions are defined as follows:

\[
\begin{align*}
\tilde{\chi}_1 &= \{y^{t-1} : H_1 y^{t-1} \leq h_i \land y(t - k) \leq V_{\tilde{\chi}}\}, \\
\tilde{\chi}_2 &= \{y^{t-1} : H_1 y^{t-1} \leq h_i \land -y(t - k) < -V_{\tilde{\chi}}\}.
\end{align*}
\]

**Step 7:** For the sub-regions \( \tilde{\chi}_1 \) and \( \tilde{\chi}_2 \), a local TS model is identified. First, we split the data belonging to the region \( \tilde{\chi}_1 \) into the two new regions, by the rule: if \( y(t - k) \leq V_{\tilde{\chi}} \) then \((y(t), (y^{t-1})^T, (u^{t-1})^T) \in \tilde{\chi}_1\), else \((y(t), (y^{t-1})^T, (u^{t-1})^T) \in \tilde{\chi}_2\). The improved significantly by the inclusion of new switching points. So, let assume we have generate a partition \( \{\tilde{\chi}_i\}_{i=1}^{s} \) and then we will analyze the inclusion of a new switching point in the model, by splitting the region \( \tilde{\chi}_i \) into two new regions divided by the new switching point. So, let consider the switching point candidate, with the maximum rate, given by:

Finally, the model parameters for the rule \( \Theta_{ij} \) in the region \( i \) can be obtained using the least-squares identification method as follows:

\[
\Theta_{ij} = \left(\psi_i^T \psi_{ij}\right)^{-1} \psi_i^T y_{ij}. \tag{16}
\]
where the matrices $\Psi_{ij}$ and $Y_{ij}$ are the following:

$$
\Psi_{ij} = \begin{bmatrix}
\beta_{ij}(z_0)(y_0)^T(u_0)^T 1 \\
\beta_{ij}(z_1)(y_1)^T(u_1)^T 1 \\
\vdots \\
\beta_{ij}(z_{N_{ij}})(y_{N_{ij}})^T(u_{N_{ij}})^T 1
\end{bmatrix},
$$

(17)

$$
Y_{ij} = \begin{bmatrix}
\beta_{ij}(z_0)y(1) \\
\beta_{ij}(z_1)y(2) \\
\vdots \\
\beta_{ij}(z_{N_{ij}}y(N_{ij}))_{N_{ij},1}
\end{bmatrix},
$$

(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 $\beta_{ij}(z(t-1)) \geq \delta$, with $\delta$ 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:

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

(19)

where $h(t)[m]$ is the level of the tank, $u(t) = Q(t)[m^3/s]$ is the input flow, outflows are $F_1(t) = S_1\sqrt{2gh(t)}$, $F_2(t) = S_2\sqrt{2gh(t) - h_1}$, $F_3(t) = S_3\sqrt{2gh(t) - h_2}$, $A_1 = 0.0154[m^2]$ 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 $y^{t+1} = h(t-1)$ as the output and $u^{t+1} = Q(t-1)$ 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 ($3$): The first one is $\mathbf{x}_{31}$, where $h(t-1) < 0.25$. The second is $\mathbf{x}_{22}$, where $h(t-1) < 0.385$ and $h(t-1) \geq 0.25$. The third is $\mathbf{x}_{23}$, where $h(t-1) \geq 0.385$. The structure of hybrid-fuzzy model is given by:

$$
R_{ij} : \text{if } h(t-1) \in \mathbf{x}_{i1} \text{ and } h(t-1) \text{ is } A_{ij1}, \text{ and } Q(t-1) \text{ is } A_{ij2}, \text{ then } h_{ij}(t) = a_{ij1}h(t-1) + b_{ij}Q(t-1) + r_{ij},
$$

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

Table 1. Parameters of hybrid-fuzzy model-2

<table>
<thead>
<tr>
<th>$\mathbf{x}_{31}$</th>
<th>$\mathbf{x}_{21}$</th>
<th>$\mathbf{x}_{22}$</th>
<th>$\mathbf{x}_{23}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{1,1}$</td>
<td>$c_{1,2}$</td>
<td>$c_{1,3}$</td>
<td>$c_{1,4}$</td>
</tr>
<tr>
<td>1</td>
<td>4.484</td>
<td>183.27</td>
<td>0.4689</td>
</tr>
<tr>
<td>2</td>
<td>5.369</td>
<td>153.32</td>
<td>0.4493</td>
</tr>
<tr>
<td>3</td>
<td>5.118</td>
<td>160.56</td>
<td>0.5224</td>
</tr>
</tbody>
</table>

$\mathbf{x}_{11}$

<table>
<thead>
<tr>
<th>$\mathbf{x}_{12}$</th>
<th>$\mathbf{x}_{13}$</th>
<th>$\mathbf{x}_{14}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{2,1}$</td>
<td>$c_{2,2}$</td>
<td>$c_{2,3}$</td>
</tr>
<tr>
<td>1</td>
<td>4.251</td>
<td>122.14</td>
</tr>
<tr>
<td>2</td>
<td>4.712</td>
<td>214.38</td>
</tr>
<tr>
<td>3</td>
<td>4.440</td>
<td>227.54</td>
</tr>
</tbody>
</table>

$\mathbf{x}_{14}$

<table>
<thead>
<tr>
<th>$\mathbf{x}_{14}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{3,1}$</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>$c_{4,1}$</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>STEPS</th>
<th>TS</th>
<th>H-F1</th>
<th>H-F2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0405</td>
<td>0.0354</td>
<td>0.0353</td>
</tr>
<tr>
<td>5</td>
<td>0.0420</td>
<td>0.0378</td>
<td>0.0363</td>
</tr>
<tr>
<td>10</td>
<td>0.0448</td>
<td>0.0395</td>
<td>0.0378</td>
</tr>
</tbody>
</table>

A switching point was detected at $h(t-1) = 0.385$. In the case of $h(t-1) = 0.385$, 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.
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 consists 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 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