Fuzzy Clustering Based Models Applied To Petroleum Processes

MIGUEL ANGEL RAMIREZ CANELON
ELIEZER COLINA MORLES
Universidad de los Andes
Facultad de Ingeniería
Complejo La Hechicera, Mérida, Venezuela
Ramirezmb@pdvsa.com
Ecolina@ula.ve

Abstract: The application of fuzzy clustering techniques has recently become in a very useful alternative in the area of modeling and identification of complex industrial processes. In particular, fuzzy clustering techniques such as Fuzzy C-Means and the Gustafson-Kessel (GK) algorithms will be analyzed and applied in details in this paper. These algorithms will be implemented in the construction of Takagi-Sugeno fuzzy models for the gas-liquid separation process, the water-oil separation process and the oil-heating process, which are important processes in the oil industry. Validations of the obtained fuzzy models will be performed and some conclusions will be established.

Key Words: Fuzzy clustering, Fuzzy C-Means, Gustafson-Kessel (GK) algorithm, least-squares method, artificial lift production methods, production separator, washing-tanks and fired heaters.

1 Introduction

Modeling and identification are essentially the first steps previous for designing control, supervision and failure detection systems. The traditional approach for modeling is based on the precise knowledge of the system under consideration and its dynamic behaviour that along with an appropriate mathematical formulation, may lead to a representative model of the system [3]. This type of traditional first-principles models are usually known as white box models or mechanistic models which are based on deep knowledge of the nature of the system. However, there are some quite complex physical processes that are not amenable to conventional modeling approaches, due to the lack of precise, formal knowledge about the system, the presence of strong nonlinearities and high degree of uncertainty and / or due to time varying characteristics [4].

A simple and intuitive technique for modeling complex processes can be achieved by dividing the global model of the system into a set of local models, where each local model has got its specific validity range. The global model of the system may be obtained by means of the integration of all the local models through the use of a fuzzy rules base that selects the appropriate local model according to the specific condition of the process [2]. The determination of the correct number of local models and the identification of these models are the major drawbacks of this sort of technique. There is an approach that contemplates the use of measured data of the process, the selection of the structure of the system based on fuzzy rules and the estimation of local parameters (parametric identification) [1]. Basically, fuzzy clustering algorithms are the most frequently used in the fuzzy identification of complex system [5][6], though there exists some other alternatives based for instance in neurofuzzy approaches [7].

In general, fuzzy identification is an effective tool for the approximation of dynamic nonlinear system on the basis of measured data [13]. Among the different fuzzy modeling techniques, the Takagi-Sugeno model has attracted most attention [14]. This model consists of IF-THEN rules with fuzzy antecedents and mathematical functions in the consequent part. The antecedent fuzzy sets partition the input space into a number of fuzzy regions, while the consequent functions describe the system's behaviour in these regions. The construction of a Takagi-Sugeno model is generally done in two steps. In the first step, the antecedent fuzzy sets are determined either by means of heuristic knowledge of the process under consideration or by using some data-driven techniques. In the second step, the consequent functions parameters are estimated. As these functions are usually chosen to be linear in their parameters, standard linear least-squares methods can be employed as estimation technique [15].

The rest of the paper is organized as follows. In section 2, the Takagi-Sugeno (TS) fuzzy model for dynamic nonlinear system is analyzed. Section 3 describes in details the Fuzzy C-Means and Gustafson-Kessel algorithms and an overview of the least squares method for the determination of the consequents of a TS fuzzy model is presented. Section 4 will briefly describe some operations related to the oil industry and three cases of study will be presented. The cases of study deal with the determination of fuzzy models for the gas-liquid separation process, the water-oil separation process and the oil-heating process on the basis of measured data; some computer simulation results will be presented and a computer validation of the fuzzy models obtained will also be performed. Conclusions are given in section 6.

2 Takagi-Sugeno Fuzzy Model For Dynamic Nonlinear System

Nonlinear dynamic systems are often represented in the Nonlinear Auto Regressive with eXogenous input (NARX) model form [16]. This kind of model establishes a nonlinear relation between the past inputs and outputs and the predicted outputs of the system, which can be represented as follows:

$$\mathbf{y}(\tau+1) = f([y_1(\tau),..., y_1(\tau-n_y),..., y_p(\tau),..., y_p(\tau-n_y), u_1(\tau),..., u_1(\tau-n_u),...u_q(\tau),..., u_q(\tau-n_u)]^T) \quad \textbf{(1)}$$

Here n_y and n_u denotes the maximum lags considered for the output and input terms, respectively. Equation (1) represents a NARX model with q inputs and p outputs and f represents the mapping of the NARX model. In every NARX model, there exists a term called state vector which is also known as the regressor vector that can be written as follows:

$$\mathbf{x}(\tau) = [y_1(\tau),..., y_1(\tau-n_y),..., y_p(\tau),..., y_p(\tau-n_y), u_1(\tau),..., u_1(\tau-n_u),...u_q(\tau),..., u_q(\tau-n_u)]^T \quad \textbf{(2)}$$

If $\mathbf{x}(\tau)$ from (2) is substituted in (1), then the following compact equation is obtained:

$$\mathbf{y}(\tau+1) = f(x(\tau)) \quad \textbf{(3)}$$

Where $\mathbf{y}(\tau+1)$ is the regressand or outputs predicted vector(MIMO case). Let's consider the identification of the nonlinear system:

$$\mathbf{y}_k = f(\mathbf{x}_k) \quad \textbf{(4)}$$

Based on some availabe input-output measured data $\mathbf{x}_k = [\mathbf{x}_{1,k} ,..., \mathbf{x}_{n,k}]^T$ and $\mathbf{y}_k$, respectively, where k = 1,2,…,N denotes the individual samples of each variable. As it is difficult to get a model that describes the system as a whole, it is always possible to construct local linear models with representation capabilities around specific selected operating points. The modeling framework that is based on combining local models valid in predefined operating regions is called Operating Regime-Based Modeling [15]. In this framework, the model is generally given by:

$$\mathbf{y}_k = \sum_{i=1}^{c} \phi_i(\mathbf{x}_k)(\mathbf{a}_i^T \mathbf{x_k} + b_i) \quad \textbf{(5)}$$

Where $\phi_i(\mathbf{x}_k)$ is the validity function for the i-th operating region and $\theta_i = [\mathbf{a}_i^T , b_i]^T$ is the parameter vector of the corresponding local linear model[16]. The operating regime can also be represented by fuzzy sets in which case the Takagi-Sugeno fuzzy model is obtained:

$$R_i: \textbf{ If } \mathbf{x}_k \textbf{ is } \mathbf{A}_i(\mathbf{x}_k) \textbf{ Then } \mathbf{y}_k = \mathbf{a}_i^T \mathbf{x}_k + b_i, \quad [w_i] \quad i = 1,...c. \quad \textbf{(6)}$$

Here $\mathbf{A}_i(x)$ is a multivariable membership function, $\mathbf{a}_i$ and b_i are the parameters of the local linear model, and $w_i \in [0,1]$ is the weight of the i-th rule. The value of w_i is usually chosen by the designer of the fuzzy system to represent the belief in the accuracy of the i-th rule. When such knowledge is not available $w_i = 1$, $\forall$i is used. The antecedent proposition "$\mathbf{x}_k$ *is* $\mathbf{A}_i(\mathbf{x_k})$" can be expressed as a logical combination of proposition with univariate fuzzy sets defined for the individual components of $\mathbf{x_k}$, generally in the following conjunctive form:

$$R_i: \textbf{If } \mathbf{x}_{1,k} \textbf{ is } \mathbf{A}_{i,1}(\mathbf{x}_{1,k}) \textbf{ and ... and } \mathbf{x}_{n,k} \textbf{ is } \mathbf{A}_{i,n}(\mathbf{x}_{n,k}) \textbf{ Then } \mathbf{y}_k = \mathbf{a}_i^T \mathbf{x} + b_i, \ [w_i] = 1,...,c \quad \textbf{(7)}$$

The degree of fulfillment of the rule is then calculated as the product of the individual membership degrees and the rule's weight:

$$\beta_i(\mathbf{x}_k) = w_i \mathbf{A_i}(\mathbf{x}_k) = w_i \prod_{j=1}^{n} A_{i,j}(\mathbf{x}_{j,k}) \quad (8)$$

The rules are aggregated by using the fuzzy-mean formula:

$$\mathbf{y}_k = \frac{\sum_{i=1}^{c} \beta_i(\mathbf{x}_k)(\mathbf{a}_i^T x + b_i)}{\sum_{i=1}^{c} \beta_i(\mathbf{x}_k)} \quad (9)$$

From (5) to (9) it can be observed that the TS fuzzy model is equivalent to the Operating Regime-Based Model when the validity function is chosen to be the normalized rule degree of fulfillment:

$$\phi_i(\mathbf{x}_k) = \frac{\beta_i(\mathbf{x}_k)}{\sum_{i=1}^{c} \beta_i(\mathbf{x}_k)} \quad (10)$$

3 Fuzzy Clustering Algorithms

Fuzzy clustering techniques are mostly unsupervised algorithms that are used to decompose a given set of objects into subgroups or clusters based on similarity. The goal is to divide the data set in such a way that objects (or example cases) belonging to the same clusters are as similar as possible, whereas objects belonging to different clusters are as dissimilar as possible. [8]. However, the assignments of objects to the classes and the description of these classes are unknown.

Fuzzy cluster analysis allows the formation of gradual membership of measured data points to clusters as degrees in the range [0, 1]. Furthermore, these membership degrees offer a much finer degree of detail of the data model. Aside from assigning a data point to clusters in shares, membership degrees can also express how ambiguously or definitely a data point should belong to a cluster. The concept of these membership degrees is substantiated by the definition and interpretation of fuzzy sets [9].

A common concept that exist in all fuzzy clustering algorithms is that they are prototype-based, i.e., the clusters are represented by clusters prototypes, v_i, $i = 1,\ldots,c$., which are used to capture the structure (distribution) of the data in each cluster. With this representation of the cluster, let us denote the set of prototypes $V = [v_1,\ldots,v_c]$. Each prototype v_i is a n-tuple of parameters that consists of a cluster center c_i and some additional parameters about the size and the shape of the cluster. The cluster center c_i is an instantiation of the attributes used to describe the domains. The size and shape parameters of a prototype determine the extension of the cluster in different directions of the underlying domain. The prototypes are constructed by the clustering algorithms and serve as prototypical representations of the data points in each cluster [9].

On the other hand, fuzzy clustering algorithms are based on objective functions J, which are mathematical criteria that quantify the goodness of cluster models that comprise prototypes and data partition. Objective functions serve as cost functions that have to be minimized to obtain optimal clusters solutions. Having defined such a criterion of optimality, the clustering task can be formulated as a function optimization problem whose algorithms determine the best decomposition of a data set into a predefined number of clusters by minimizing its objective function [8].

Fuzzy clustering techniques can be applied to data that are quantitative (numerical), qualitative, or a mixture of both and in general, such data are related to observed data of some specific process [10]. The available data samples are collected in matrix **Z**, formed by concatenating the regression data matrix **x** and the output vector **y:**

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{x}_n^T \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{y}_n \end{bmatrix}, \quad \mathbf{Z^T} = [\mathbf{x}, \mathbf{y}] \quad (11)$$

Thus, each observation is a n+1 dimensional column vector expressed as: $\mathbf{z}_k = [\mathbf{x}_{1,k}, \ldots, \mathbf{x}_{n,k}, \mathbf{y}_k]^T = [\mathbf{x}_k, \mathbf{y}_k]^T$. The columns of Z are denoted by z_l, $l = 1,\ldots,N$, Where N is the number of points used in the identification of the system.

3.1 Fuzzy C-Means Objective Function

The most common objective function used in fuzzy clustering algorithms is given by:

Where:

$$J(Z;U,V) = \sum_{i=1}^{c} \sum_{k=1}^{N} (\mu_{ik})^m \|z_k - v_i\|_B^2 \quad (12)$$

$$U = [\mu_{ik}] \in [0,1],\ 1 \le i \le c,\ 1 \le k \le N \qquad \textbf{(13)}$$

is a fuzzy partition matrix of Z,

$$V = [v_1, v_2, ..., v_c],\ \ v_i \in \Re^n \qquad \textbf{(14)}$$

is a vector of cluster prototypes(centers) to be determined,

$$D^2_{ik\,B} = \|z_k - v_i\|^2_B = (z_k - v_i)^T B (z_k - v_i) \quad \textbf{(15)}$$

is a norm that is determined by the selection of the matrix B. In case matrix B is equal to the identity matrix (B = I), the norm becomes in the Euclidean distance and equation (15) can be written as:

$$D^2_{ik} = (z_k - c_i)^T (z_k - c_i) \qquad \textbf{(16)}$$

This particular option implies the generation of hyper-spherical classes as depicted in figure 1a. However, B can be chosen as a diagonal matrix nxn, where n, is the dimension of the elements of Z whose values may take different variances in each of the directions of the coordinated axis of Z (a covariance matrix):

$$B = \begin{bmatrix} (1/\sigma_1^2) & 0 & ... & 0 \\ 0 & (1/\sigma_2^2) & ... & 0 \\ . & . & & . \\ . & . & ... & . \\ 0 & 0 & ... & (1/\sigma_n^2) \end{bmatrix} \qquad \textbf{(17)}$$

Thereby, hyper-elliptical clusters as show in figure 1b, are created. Another option is to define B as the inverse of the covariance matrix of the samples of Z. For this particular case, clusters depicted in figure 1c are generated. Finally:

$$m \in (1, \infty) \qquad \textbf{(18)}$$

is a parameter which determines the fuzziness of the resulting clusters. In general, the value of the cost function (12) can be seen as a measure of the total variance of z_k from v_i [1].

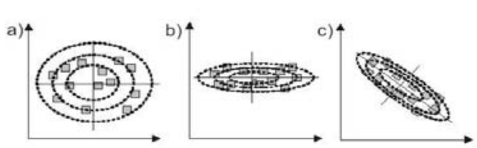

Figure 1: Norms used in fuzzy clustering

3.2 Fuzzy C Means Algorithm

The minimization of the c-means functional (12) represents a nonlinear optimization problem that can be solved by using a variety of methods, including iterative minimization. The most popular method is a simple Picard iteration through the first order conditions for stationary points of (12), known as the fuzzy c-means algorithms [10]. Let's define the next constraints:

$$\forall k \ \sum_{i=1}^{c} \mu_{ik} > 0 \quad \textbf{(19)}$$

$$\forall i \ \sum_{k=1}^{N} \mu_{ik} = 1 \quad \textbf{(20)}$$

Constraint (19) guarantees that no cluster is empty and constraint (20) ensures that the sum of the membership degrees for each data set is equal to 1 In other words, each data set receives the same weight in comparison to all other data and therefore all data set are equally included into the cluster partition.

The stationary points of the objective function (12) can be found by adjoining the constraint (18) to J by means of Lagrange Multipliers:

$$J(Z;U,V,\lambda) = \sum_{i=1}^{c}\sum_{k=1}^{N} (\mu_{ik})^m D^2_{ik\,\mathbf{A}} + \sum_{k=1}^{N} \lambda_k \left(\sum_{i=1}^{c} \mu_{ik} - 1 \right) \quad \textbf{(21)}$$

And by setting the gradients of J with respect to U, V y λ to zero, it can be shown [18] that the stationary points correspond to:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{D^2_{ik\,\mathbf{A}}}{D^2_{jk\,\mathbf{A}}} \right)^{\frac{2}{m-1}}};\ 1 \le i \le c \quad \textbf{and} \quad 1 \le k \le N \quad \textbf{(22)}$$

$$v_i = \frac{\sum_{k=1}^{N} (\mu_{ik})^2 z_k}{\sum_{k=1}^{N} (\mu_{ik})^2};\ 1 \le i \le c \qquad \textbf{(23)}$$

Equation (23) gives a value for v_i that can be interpreted as the weighted average that belongs to a cluster and whose weights are given by the term μ_{ik}. The major problem that arises when using this kind of algorithm for the identification of fuzzy models is that the identified clusters have hyper-ellipsoid forms. For a future control application, it is not convenient to have this sort of cluster, instead, it will be adequate to get clusters with some kind of linear structure.[1]. There exist some other approaches applied to the fuzzy c means technique. These approaches consist in the use of an adaptive distance, where each norm is different for each cluster; thus it would be possible to get clusters of data with different structures.

3.3 Gustafson-Kessel Algorithm

Gustafson and Kessel proposed a powerful fuzzy clustering algorithm (GK) that is based on adaptive distance measures. It can be used to obtain a nonlinear model from a collection of local linear models. This technique has a model structure that is easy to understand and interpret, and can integrate various types of knowledge, such as empirical knowledge, derived from first-principles and measured data [11].

This algorithm has been frequently used because most of the time the clusters obtained have quasi-linear behaviour at different operating regimes that may exist in the data set and therefore, it is ideal for constructing Takagi-Sugeno Fuzzy Models [14].

Thus, if a different norm B_i is selected for each cluster and if equation (15) is taken into account then:

$$D^2_{ik\,B_i} = \|z_k - v_i\|^2_{B_i} = (z_k - v_i)^T B_i (z_k - v_i) \quad \textbf{(24)}$$

Now, these matrices are used as the variables for the optimization of function (12) adapting the specific norm for each cluster according to its characteristics. Let's define $B_i = \{B_1, B_2, \ldots, B_C\}$ as the vector that contains the v norms. The new function to be minimized taking into account (13), (14) and (18) is:

$$J(Z;U,V,B) = \sum_{i=1}^{c}\sum_{k=1}^{N} (\mu_{ik})^m \|z_k - v_i\|^2_{B_i} \quad \textbf{(25)}$$

This objective function can not be directly minimized with respect to B_i , since it is linear with respect to B_i .To obtain a feasible solution, B_i must be constrained in some way. The usual way of accomplishing this is by constraining the determinant of B_i :

$$|B_i| = \rho_i, \quad \rho > 0 \quad \textbf{(26)}$$

With ρ_i a constant for each cluster. Once the Lagrange multiplier is used, the next solution is obtained:

$$|B_i| = [\rho_i \det(F_i)]^{\frac{1}{n}} F_i^{-1} \quad \textbf{(27)}$$

Where n is the dimension of the element of Z and F_i is the fuzzy covariance matrix of the i-th cluster given by:

$$F_i = \frac{\sum_{k=1}^{N} (\mu_{ik})^m (z_k - v_i)(z_k - v_i)^T}{\sum_{k=1}^{N} (\mu_{ik})^m} \quad \textbf{(28)}$$

When all the clusters are obtained, the determination of a Takagi-Sugeno model can be achieved.

Summarize of the GK Algorithm: Given the data set Z, choose the weighting exponent $m>1$(the most common value is m = 2.2), the termination tolerant ε (it is recommended to select a value in the range [0.001; 10^{-6}]), the number of fuzzy rules or clusters $c \in [1,C]$, N is the number points for the identification system and l is the number of iteration. Initialize the fuzzy partition matrix randomly $U^{(l)}$.

Repeat for l=1,2,…
Step 1: Compute cluster means (prototypes).

$$v_i^{(l)} = \frac{\sum_{k=1}^{N} \left(\mu_{ik}^{(l-1)}\right)^m z_k}{\sum_{k=1}^{N} \left(\mu_{ik}^{(l-1)}\right)^m}, \quad 1 \le i \le C$$

Step 2: Compute covariance matrices.

$$F_i = \frac{\sum_{k=1}^{N} (\mu_{ik})^m (z_k - v_i^{(l)})(z_k - v_i^{(l)})^T}{\sum_{k=1}^{N_d} (\mu_{ik})^m}, \quad 1 \le i \le c$$

Step 3: Compute distances.

$$D^2_{ik} B = \left(z_k - v_c^{(l)}\right)^T [\rho_i \, det(F_i)]^{\frac{1}{n}} F_i^{-1} \left(z_k - v_c^{(l)}\right)$$
$$1 \le i \le c \quad \textbf{and} \quad 1 \le k \le N$$

Step 4: Update of the fuzzy partition matrix.
If $D_{ik}B > 0 \ \ \forall$ i[1,C] and k[1,N].

$$\mu_{ik}^{(l)} = \frac{1}{\sum_{j=1}^{c}\left(\frac{D_{ikB}}{D_{jkB}}\right)^{\frac{2}{m-1}}}$$

Otherwise:

$$\mu_{ik}^{(l)} = 0, \textbf{If } D_{ik}B > 0 \textbf{ and } \ \mu_{ik}^{(l)} \in [\mathbf{0,1}] \textbf{ with} \sum_{\mathbf{i=1}}^{\mathbf{c}} \mu_{ik}^{(l)} = 1$$

Step 5: Evaluation of ε for l iteration (l >= 2).

$$\text{If} \quad \left|U^{(l)} - U^{(l-1)}\right| < \varepsilon, \quad \text{end of the algorithm.}$$

Otherwise, go to Step one and $l = l + 1$.

3.4 Least-Squares Method.

The antecedent membership functions A^k can be computed, and hence the consequent parameters $\mathbf{a}_i^T$ and b_i can be calculated using the least-squares method [17].

For the determination of the antecedent membership functions, it is important to consider that each cluster obtained is a Takagi-Sugeno fuzzy rule. The multidimensional membership function A^k are given analytically by computing the distance of $\mathbf{x}(\tau)$ (Regressor Vector) from the projection of the cluster prototypes center v_i onto X, and then computing the membership degree in an inverse proportion to the distance. Denote with F_x , a submatrix of F_i, which, describes the form of the cluster in the antecedent space X. Let $V_x = [v_1, v_2,\ldots,v_k]^T$ denote the projection of the cluster center onto the antecedent space X. From the Gustafson-Kessel algorithm, the inner product distance norm, given by:

$$D_{kl} = (x(\tau) - V_x)^T \left|F_x\right|^{1/n} (F_x)^{-1} (x(\tau) - V_x) \quad \textbf{(29)}$$

is converted into the membership degree by:

$$\mu_{A^k}(\mathbf{x}(\tau)) = \frac{1}{\sum_{j=1}^{K}\left(d_{kl}/d_{jl}\right)^{2/(m-1)}} \qquad \textbf{(30)}$$

For the determination of the consequent parameters, as it was pointed out before, least-squares methods can be used. Let $\Gamma_k \in \Re^{NxN}$ denote the diagonal matrix having the membership degree $\mu_{Ak}x(\tau)$ as its k-th diagonal element. By appending a unitary column to **X**, the extended matrix X_e = [**X**,1] is created. Furthermore, denote X', the matrix composed of the products of matrices Γ_i y X_e:

$$X' = [\Gamma_1 X_e, \Gamma_2 X_e, \ldots, \Gamma_k X_e] \quad \textbf{(31)}$$

The consequent parameters $\mathbf{a}_i^T$ and b_i are lumped into a single parameter vector:

$$\theta = [\mathbf{a}_1^T, b_1, \mathbf{a}_2^T, b_2, \ldots, \mathbf{a}_k^T, b_k] \quad \textbf{(32)}$$

Given the data **X** and **y,** it is well known that the solution of y = X'θ + ε, by means of least-squares techniques has the solution:

$$\theta = [(X')^{-1} X']^{\mathbf{T}} (X')^T \mathbf{y} \quad \textbf{(33)}$$

This is an optimal least-squares solution which gives the minimal prediction error. Nevertheless, a weighted least-squares approach can be applied and then the next solution is obtained:

$$[\mathbf{a}_k^T, b_k]^T = [X_e^T \, \Gamma_k \, X_e]^{\mathbf{-1}} X_e^T \Gamma_k \mathbf{y} \quad \textbf{(34)}$$

4. Major Operations in oil processes

Oil processes constitute a set of fundamental operations, which in general may be divided in three phases: extraction, processing and transport. In the process of extraction, since some petroleum's volume has low gravity and due to some specific characteristic of some wells (heavy oil), it is necessary the installation of some artificial lift methods in order to extract the petroleum from the oilfields. Among these artificial lift methods, the most common are: mechanical pumping, progressive cavity pumping and electro submersible pumping. In the extraction phase, the petroleum (crude) in its original state (gas, water and petroleum) is pumped to a specific place called discharge station. The function of a discharge station is to separate the oil well stream into three components (oil, gas and water) and to process these phases into some marketable products or dispose of them in an environmentally acceptable manner. In a discharge station, the sum of the petroleum's flow of a determined group of wells is received by a production multiple. In the production multiple, the petroleum stream is distributed to a group of production separators.

Figure 2 shows the drawing of a standard production separator. A separator for petroleum production is a large drum designed to separate production fluids into their constitute components of oil, gas and water. It operates on the principle that the three components have different densities, which allows them to stratify when moving slowly with gas on top, water on the bottom and oil in the middle.

A standard separator has two liquid outputs and one gas output. In the oil output, there is a control valve, which is used to regulate the liquid of the separator; the water output is used only for maintenance purpose. The gas-output flows toward a system that keeps some other instruments operating, for instance, all the actuators of the station. Also, the gas is used as a combustible material for keeping the furnaces operative.

The liquid component at the output of the separator flows toward the fired heaters, where, the temperature of the liquid is increased from 95 degrees Fahrenheit to 190 degrees Fahrenheit, thus reducing the viscosity of the liquid. Once the temperature of the liquid is increased, the liquid flows toward some special tanks (Washing-Tanks) where the oil and water are completely separated. The oil is pumped to a major station either for being refined or for marketing purposes.

4.1 First case of study: gas-liquid separation process

As it was mentioned before, the basic function of a separator is to separate the gas and liquid phases. Because of the significant variations in the productions of some wells, the flow of petroleum at the input of the separator has a high nonlinear behaviour. Currently, control systems are based on PID algorithms, with a rather poor performance.

Figure 2 illustrates the simplified separation process, where an oil-well fluid with molar F_{in} and gas, oil and water molar fractions Z_g, Z_o, Z_w respectively enters the separator. The hydrocarbon component of the fluid separates into two parts; the first stream F_{h1} separates by gravity and enters the oil phase, and the second stream F_{h2} stays in the aqueous phase due to incomplete separation. The liquid discharge from the aqueous phase Fw_{out} is a combination of the dumped water stream Fw plus the unseparated hydrocarbon stream F_{h2}. The gas component in the separated hydrocarbon stream, which enters the oil phase, separates into two parts: the first gas stream F_{g1} flashes out of the oil phase due to the pressure drop in the separator and the second gas stream F_{g2} stays dissolved in the oil phase. The oil discharge Fo_{out} of the separator contains the oil component of the separated hydrocarbon Fo and the dissolved gas component F_{g2}. The flashed gas Fg_{out} flows out of the separator for further processing [20].

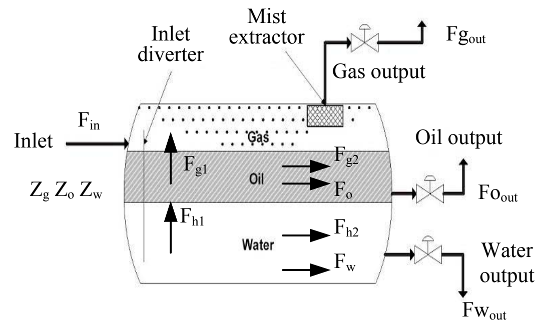

Figure 2: Drawing of a production separator

Based on the fact, that a fuzzy model-based control will be implemented in the future, a fuzzy model will be constructed based on historical data. A separator has a series of instruments installed for measuring certain variables: the pressure of the separator, the temperature of the liquid, the level of the separator and a valve control with a pneumatic actuator. Also, there are switches for detecting low level and high level of petroleum and high pressure in the separator.

A set of 1000 sampled data per variable was recorded with a time sample of one second. It was taken into account three operation regimes: level of liquid greater than the level of reference, level of liquid near the level of reference and level of liquid less than the level of reference.

The selected input variables were: Current Level of the Separator $y(\tau)$ and Valve Position $u(\tau)$. The output variable was the level at the next sampling instant $y(\tau+1)$. The model consisted of three rules with linear consequents, including the bias term. The rule base represents a non-linear first-order regression model that can be written as follows:

$$y(\tau+1) = f(y(\tau), u(\tau)) \qquad \textbf{(35)}$$

The Fuzzy C-Means algorithm was applied to the set of 1000 data sampled. Figure 3 shows the clusters obtained. The same procedure was applied for the Gustafson-Kessel algorithm. Figure 4 shows the clusters obtained. The MATLAB Software was used to carry out the simulations.

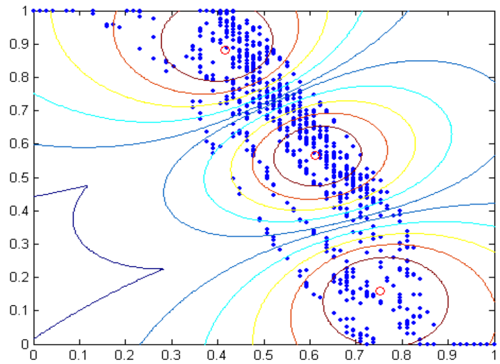

Figure 3: Application of fuzzy c-means algorithm based on historical data of the production separator

The index fuzziness (weighting exponent) *m* was set to 2.2, C (Number of total clusters) was set to 3 and ε was set to 10^{-6}. The VAF variance accounted performance index obtained was 97,80 %, so it is a good indication of the precision of the model obtained. The number of iterations *l* obtained was 123 for the GK algorithm and *l* = 156 for the FCM algorithm.

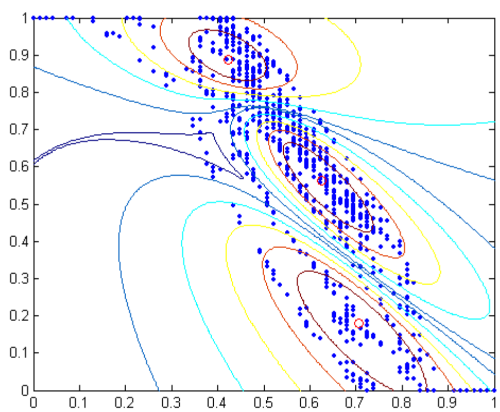

Figure 4: Application of Gustafson-Kessel algorithm based on historical data of the production separator

For the construction of the Takagi-Sugeno fuzzy model, first the antecedent membership function were determined by projecting the fuzzy partition matrix onto the two antecedent variable as it is shown generically in figure 5. Basically, the equations (29) and (30) are used for obtaining the membership functions. Figure 6 shows the resultant membership functions obtained from the simulation.

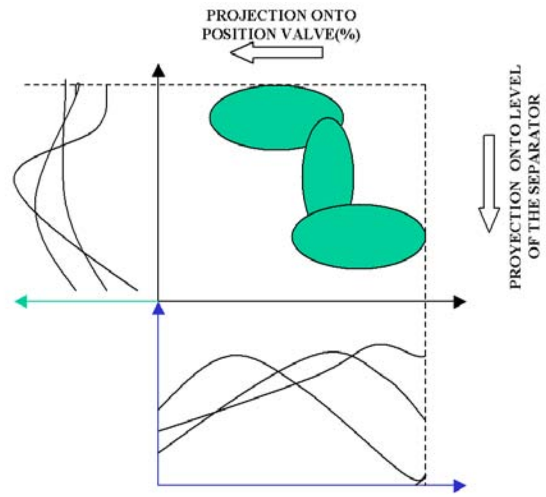

Figure 5: Projection of fuzzy clusters onto the antecedent variables

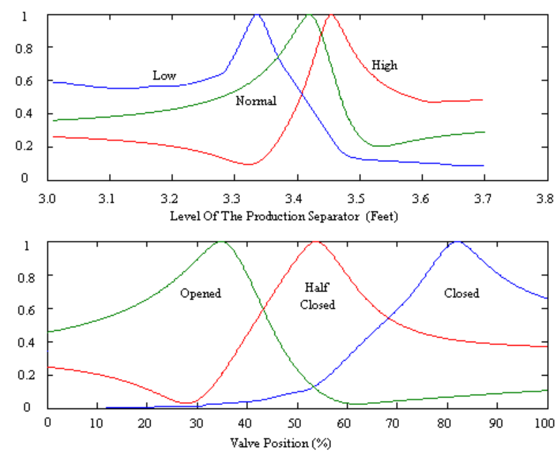

Figure 6: Membership functions obtained from the simulations

The consequent parameters were estimated by least-squares method such as the solution proposed in (33). The resultant rule base obtained from the simulation was:

1. **If** $y(\tau)$ is Low **AND** $u(\tau)$ is Opened **Then** $y(\tau+1) = 1.0697y(\tau) + 0.0002u(\tau) - 0.2491$
2. **If** $y(\tau)$ is Normal **AND** $u(\tau)$ is Half-Closed **Then** $y(\tau+1) = 0.9375y(\tau) + 0.2170$
3. **If** $y(\tau)$ is High **AND** $u(\tau)$ is Closed **Then** $y(\tau+1) = 1.0540y(\tau) + 0.003u(\tau) - 0.1978$

This fuzzy model was validated with 775 data sampled different from the set of 1000 data sample. Figure 7 shows the validation of the model.

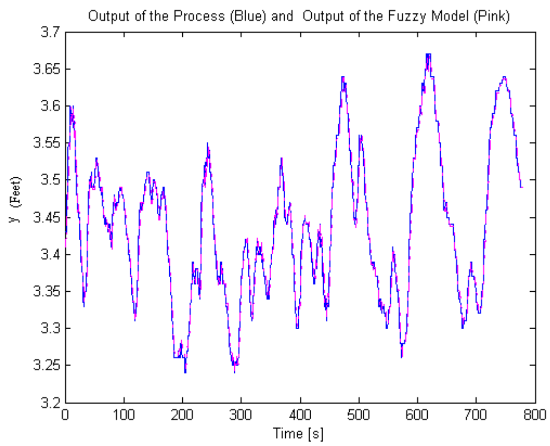

Figure 7: Validation of the fuzzy model obtained for the production separator

4.2 Second case of study: water-oil separation process

Separating the oil from the water is one of the most complex processes in the oil industry. Standard washing tanks (Figure 8) are used for this purpose. Washing tanks are passive, physical separation systems designed for removing oil from water, where, the oil-water mixture enters the tank and is spread out horizontally, distributed through an energy and turbulence diffusing device.

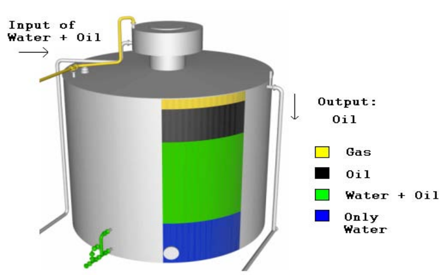

Figure 8: Illustration of a washing tank

A standard washing tank has a series of instruments installed for measuring certain variables. These variables are: the level of water inside the tank, the temperature in the input of the tank, the water-cut in the output of the tank and a valve control with a pneumatic actuator for the regulation of water level. Also, there are switches for detecting low level of water and high level of oil.

Currently, control systems are based on PID algorithm. The performance can be considered poor because of some no linearity associated with the time of separation between the molecules of oil and gas. Based on the fact that a fuzzy model-based control will be implemented, a fuzzy model of the washing tank will be constructed base on historical data.

The selected input variables were: the level of water inside the tank $y(\tau)$ and the valve position $u(\tau)$ (Input variables). The output variable was the level of water at the next sampling instant $y(\tau+1)$. Given the fact, that this sort of process is quite slow, some tests were done in order to determine the appropriate time sample. A value of 20 second was considered as an appropriate time sample.

The resultant fuzzy model was similar to the one obtained for the production separator: three rules with linear consequents including the bias term. The FCM and the GK algorithms applied to this particular process are showed in figure 9 and 10. The MATLAB Software was used to carry out the simulations. The index fuzziness (weighting exponent) m was set to 2.2, C (Number of total clusters) was set to 3 and ε was set to 10^{-4}. The VAF variance accounted performance index obtained was 90,30 %, so it is an acceptable indication of the precision of the model obtained. The number of iterations l obtained was 223 for the GK algorithm and $l = 276$ for the FCM algorithm. It is important to highlight that for $\varepsilon < 10^{-4}$, the VAF obtained was less than 87%.

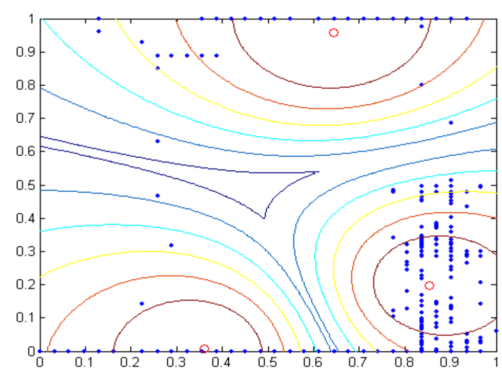

Figure 9: Application of fuzzy c-means algorithm based on historical data for the washing tank

The determination of the Takagi-Sugeno fuzzy model was derived through computer simulations. The resultant rule base obtained from the simulation was:

1. **If** $y(\tau)$ is Low **AND** $u(\tau)$ is Closed **Then** $y(\tau+1) = y(\tau) + 7{,}831u(\tau) + 0.0005$
2. **If** $y(\tau)$ is Normal **AND** $u(\tau)$ is Half-Opened **Then** $y(\tau+1) = 0.9375y(\tau) + 0.0005$

3. **If** *y(τ)* is High **AND** *u(τ)* is Opened **Then** $y(\tau+1) = 0{,}7y(\tau) + 4{,}7$

This fuzzy model was validated with 229 data sampled different from the set of 628 data sample. Figure 11 shows the validation of the model.

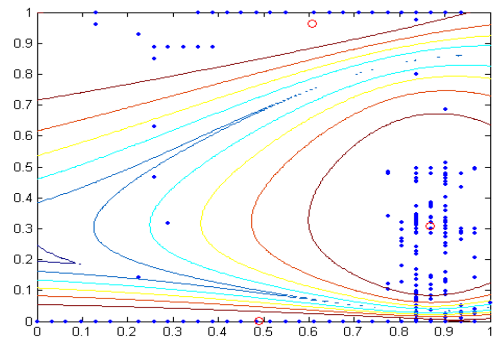

Figure 10: Application of Gustafson-Kessel algorithm based on historical data of the washing tank

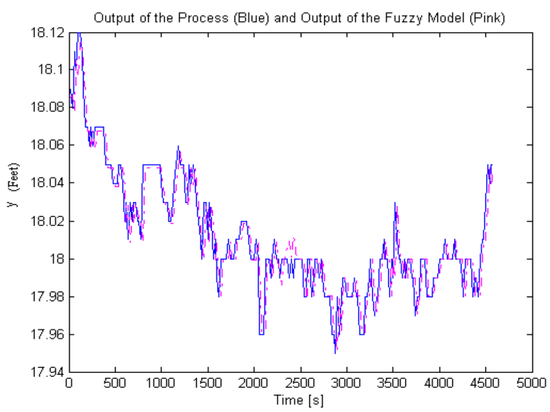

Figure 11: Validation of the fuzzy model obtained for the washing tank

The resultant membership functions are showed in figure 12:

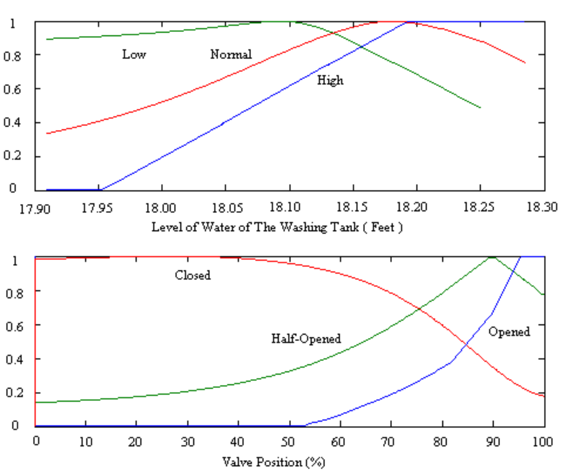

Figure 12: Membership functions obtained from the simulations

4.3 Third case of study: The oil-heating process.

A fired heater (Figure 13) is a direct-fired heat exchanger that uses the hot gases of combustion to raise the temperature of a feed flowing through coils of tubes aligned throughout the heater. Depending on the use, these are also called furnaces or process heaters [21]. Some heaters simply deliver the feed at a predetermined temperature to the next stage of the reaction process; others perform reactions on the feed while it travels through the tubes [19].

Fired heaters are used throughout hydrocarbon and chemical processing industries such as refineries, gas plants, petrochemical, chemicals and synthetics, olefins, ammonia and fertilizer plants.

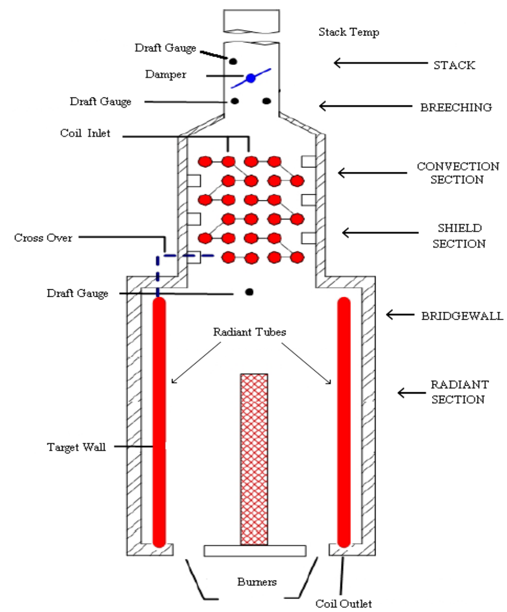

Figure 13: Illustration of a standard fired heater

Figure 13 shows a natural draft process heater with inspiration type burners. Fuel to the burners is regulated from exit feed temperature and the level of production desired determines firing rate. The most important part are the radiant section, the convection section, the shield section and the stack and breeching.

- Radiant section: the radiant tubes (horizontal or vertical) are located along the walls in the radiant section of the heater and receive radiant heat directly from the burners or target wall. The radiant zone with its refractory lining is the costliest part of the heater and most of the heat is gained there. This is also called the firebox.

- Convection Section: the feed charge enters the coil inlet in the convection section where it is preheated before transferring to the radiant tubes. The convection section removes heat from the flue gas to preheat the contents of the tubes and significantly reduces the temperature of the flue gas exiting the stack.

- Shield section: just below the convection section is the shield section, containing rows of tubing which shield the convection tubes from direct radiant heat. Several important measurements are normally made just below it. For instance, measurement of the draft at this point is important since this determines how well the heater is setup. This is also the ideal place for flue gas oxygen and PPM (parts per million) combustibles measurement.

- Stack and Breeching: the transition from convection section to the stack is called the breeching. By the time the flue gas exits the stack, most of the heat should be recovered and the temperature decreases.

A standard fired heater has a series of instruments installed for measuring a group of variables. The most important variables are: the temperature of the oil at the output, the pressure in the input and output of the fired heater, the temperature in the flue stack, the temperature in the radiant section and the temperature in the convection section. Also, there are switches that detect the state of each burner. (Normally in this sort of fired heater, there are nine burners).

The selected input variables were: the current temperature T(τ), the valve position u(τ) and error signal e(τ), which is the difference between the setpoint and the current temperature of the oil at the output of the fired heater. The output variable was the temperature of the oil at the next sampling instant T(τ+1), so the rube base may be written as follows:

$$T(\tau+1) = f(T(\tau), u(\tau), e(\tau)) \quad \textbf{(36)}$$

The Fuzzy C-Means algorithm was applied to a set of 2500 data sampled with a time sample of 12 seconds. Figure 14 shows the clusters obtained. The same procedure was applied for the Gustafson-Kessel algorithm. Figure 15 shows the clusters obtained. The index fuzziness (weighting exponent) *m* was set to 2.2. , C (Number of total clusters) was set to 6 and ε was set to 10^{-6}. The VAF variance accounted performance index obtained was 96,30 %, so it is a good indication of the precision of the model obtained. The number of iterations *l* obtained was 176 for the GK algorithm and *l* = 204 for the FCM algorithm.

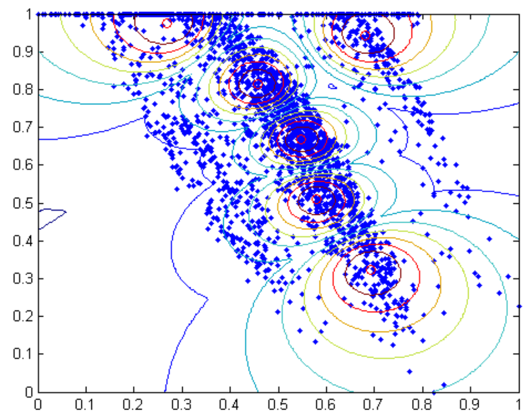

Figure 14: Application of fuzzy c-means algorithm based on historical data of the fired heater

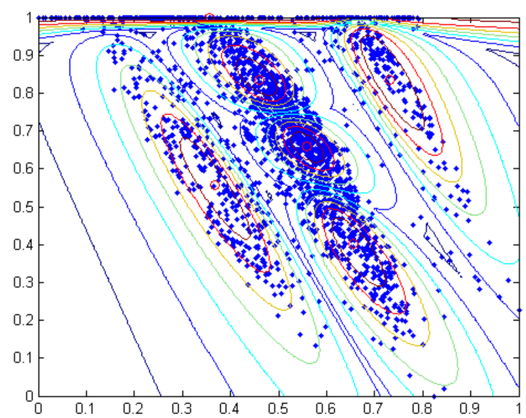

Figure 15: Application of Gustafson-Kessel algorithm based on historical data of the fired heater

The determination of the Takagi-Sugeno fuzzy (See figure 16) model was done through computer simulations. The resultant rule base obtained from the simulation was:

1. **If** *T(τ)* is Low **AND** *u(τ)* is Closed **AND** *e*(*τ*) is Negative **Then** *T(τ+1)* = 1,0620*T(τ)* +0,1291*u(τ)* +0,2781*e(τ)* –24,74.

2. **If** *T(τ)* is Low **AND** *u(τ)* is Closed **AND** *e*(*τ*) is Negative **Then** *T(τ+1)* = 0,8206*T(τ)* +0,0346*u(τ)* –0,1452*e(τ)* +29,6766.

3. **If** *T(τ)* is Normal **AND** *u(τ)* is Half Opened **AND** *e*(*τ*) is Null **Then** *T(τ+1)* = 1,0067*(τ)* +0,1601*u(τ)* -0,0865*e(τ)* –13,1905

4. **If** $T(\tau)$ is Normal **AND** $u(\tau)$ is Half Opened **AND** $e(\tau)$ is Null **Then** $T(\tau+1) = 1{,}0967(\tau) + 0{,}3601u(\tau) - 0.005e(\tau) + 11{,}2100$.

5. **If** $T(\tau)$ is Normal **AND** $u(\tau)$ is Opened **AND** $e(\tau)$ is Positive **Then** $T(\tau+1) = 2{,}0300(\tau) + 0{,}301u(\tau) - 0{,}8765e(\tau) + 18{,}2198$.

6. **If** $T(\tau)$ is High **AND** $u(\tau)$ is Opened **AND** $e(\tau)$ is Positive **Then** $T(\tau+1) = 3{,}345(\tau) + 1{,}456u(\tau) + 0.9876e(\tau) - 21{,}1905$

This fuzzy model was validated with 1000 data sampled different from the set of 2500 data sample. Figure 16 shows the validation of the model.

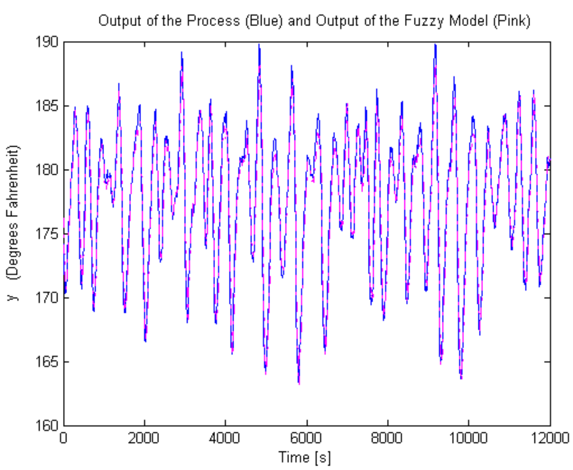

Figure 16: Validation of the fuzzy model obtained for the fired heater

The resultant membership functions are showed in figure 17:

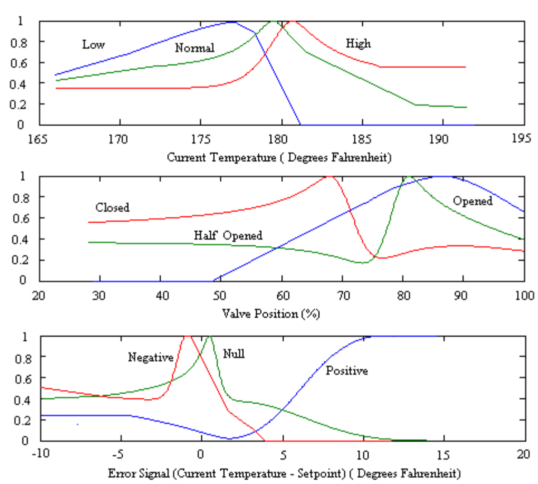

Figure 17: Membership functions obtained from the simulations

5. Conclusions

The Fuzzy C-Means and the Gustafson-Kessel algorithms were analyzed and implemented in details in this paper. Three cases of study related to the gas-liquid separation process, the water-oil separation process and a fired heater were presented for the implementation of these techniques.

All models were constructed based on a MISO framework. Two inputs and one output parameters were defined for the gas-liquid separation process and the water-oil separation process as elements of the fuzzy model with a set of 1000 sampled data for the production separator and 628 sampled data for the washing tank. A set of 2500 data sampled was used to determine the Takagi-Sugeno fuzzy model of the fired heater.

The fuzzy models were computationally validated with a set of different data samples (775 different data sample for the production separator, 229 different data sample for the washing tank and 1000 different data sample for the fired heater). All models showed an excellent generalization. This kind of modeling technique applied to complex processes gives good results in a relatively short period of time.

Finally, it is important to highlight that one of the most important advantages of this sort of modeling technique is the flexibility of incorporate qualitative information by the designer of the model along with the use of quantitative information.

References:

[1] Diez J.L., Navarro J.L, Salas A. (2004). "Algoritmos de Agrupamiento en la Identificación de Modelos Borrosos". Departamento de Ingeniería de Sistemas y Automática, Universidad Politécnica de Valencia, España, pp-1-6..

[2] Johansen T.A, Murray-Smith, R. (1997). "The Operating Regime Approach to Nonlinear Modeling and Control". Multiple Model Approaches to Modeling and Control. Taylor & Francis, London.

[3] Carbonell P., Diez J.L., Navarro J.L. (2004). "Aplicaciones de Técnicas de Modelos Locales en Sistemas Complejos". Departamento de Ingeniería de Sistemas y Automática, Universidad Politécnica de Valencia, España, pp.1-9.

[4] Babuska, R. (1998). "Fuzzy Modeling and Identification Toolbox", For use with MATLAB, pp3-5.

[6] Sugeno M., Yasukawa T. (1993). "A Fuzzy-Logic Based Approach to Qualitative Modeling". Transaction on Fuzzy Systems, Vol 1, number 1, pp 7-10.

[7] Babuska, R. (2004). "Modeling and Identification", PhD dissertation. Delft University of Technology, Delft, the Netherlands

[8] Nelles, O. (1999). "Nonlinear System Identification with Local Linear Neuro-Fuzzy Models". PhD dissertation. Darmstadt University of Technology, Darmstadt, Germany,pp 1-8.

[9]Kruse,R., Doring C., Lesot M. (2007). "Fundamental of Fuzzy Clustering", Department of Knowledge Processing and Language Engineering, University of Magdeburg, Germany, pp. 1-10.

[10] Babuska, R. (2004). "Fuzzy and Neural Control DISC Course Lecture Notes". Delft Center For Systems and Control, The Netherlands.

[11]Sousa J., Kaymak U. (2002). "Fuzzy Decision Making in Modeling and Control". World Scientific Series in Robotics and Intelligent Systems, pp. 93-108.

[12]Gustafson, D. E., Kessel W.C. (1979). "Fuzzy Clustering with a Fuzzy Covariance Matrix". Proc. IEEE CDC, San Diego, USA, pp. 761-766.

[13] Hellendoorn H., Driankov D. (1997) "Fuzzy Model Identification". Selected Approaches. Springer, Berlin, Germany, pp. 5-8.

[14] Takagi T., Sugeno M. (1985). "Fuzzy Identification Of System And Its Application To Modeling And Control". IEEE Transactions on Systems, Man and Cybernetics, pp. 116-120.

[15] Abonyi J., Babuska R., Szeifert F. (2004). "Modified Gath-Geva Fuzzy Clustering for Identification of Takagi-Sugeno Fuzzy Models". Delft University of Technology, Department of Information Technology and Systems, The Netherlands pp. 1-6.

[16] Abonyi J., Choban T., Szeifert F.(2000) "Identification of Nonlinear Systems using Gaussian Mixture of Local Models". University of Vezprem, Department of Process Engineering, Hungary, pp 3-8.

[17] Diez J.L., Navarro J.L. (1999). "Fuzzy Models of Complex Systems by Means of Clustering Techniques". Proc. 2nd Intelligent Systems in Control and Measurement, pp. 147-153.

[18] Besdek J.C.(1980). "A convergence theorem for the fuzzy ISODATA clustering algorithms". IEEE transactions, Machine Intelligent. PAMI-2(1), pp.1-8.

[19] Lobo W., Evans J.(2004). “Heat Transfer In The Radiant Section of Petroleum Heaters”. SPE transactions, pp. 1-2.

[20] Atalla F and Taylor J. (2004) “ Modeling and Control of Three-Phase Gravity Separators in Oil Production Facilities”. Department of Electrical & Computer Engineering, University of New Brunswick, Canada, pp. 1-6.

[21] Wildy, F. (2003) “ Fired Heater Optimisation”. AMETEK, Process Instrument, Pittsburgh, U.S.A, pp.1-4.