ON-LINE EXTRACTION OF FUZZY RULES IN A WASTEWATER TREATMENT PLANT

J. Victor Ramos[1,3], C. Gonçalves[2] and A. Dourado[3]

[1]*School of Technology and Management, Polytechnic Institute of Leiria; Morro do Lena – Alto do Vieiro, 2411-901 Leiria, Portugal;* [2]*Soporcel – Sociedade Portuguesa de Papel, S.A., Apartado 5 – Lavos, 3081-851 Figueira da Foz, Portugal;* [3]*Adaptive Computation Group – CISUC, Departamento de Engenharia Informática; Pólo II–Universidade de Coimbra, 3030-290 Coimbra, Portugal*

Abstract: Fuzzy systems have an important role in knowledge extraction from the huge amount of data acquired by the industrial distributed computer control systems. The paper presents a work concerning the building of a computational model of the WasteWater Treatment Plant (WWTP) in Soporcel mill (pulp and paper). Clustering of data is developed and from clusters a set of fuzzy rules describing the process behaviour is obtained, building up simple and applicable models with reasonable accuracy. Due to the time-varying dynamics of the process, on-line learning algorithms are necessary. Evolving Takagi-Sugeno (eTS) fuzzy models are used to predict the pH values in the plant The approach is based on an on-line learning algorithm that recursively develops the model structure and parameters during the operation of the process. Results for the second stage of the effluent neutralization process are presented and, despite the complexity, non-linear characteristics and time-varying dynamics of the process, the results show that the eTS fuzzy models are computationally very efficient and have practical relevance.

Key words: On-line learning; TS fuzzy models; eTS fuzzy models; subtractive clustering algorithm; recursive fuzzy clustering; pulp and paper industry.

1. INTRODUCTION

Modern information technology allows the collection, store and analysis of huge amounts of data that can be used to construct models. Each day more decisions are made based on the analysis of data; however exploiting the information contained in database systems in an intelligent way turns out to be

fairly difficult since there is a lack of tools to transform data into useful information and knowledge.

Several methods and tools are used in intelligent data analysis, including soft computing techniques. Fuzzy systems can be used to create accurate predictive models from data but it is also important that these models should be interpretable and therefore useful to describe in words the underlying data in order to support efficiently the human decision process (Setnes, 2001).

Construction of a computational model for the WWTP presents several difficulties and in the absence of a first-principles model, fuzzy modelling, and particularly Takagi-Sugeno (TS) fuzzy models, seems to be the most appropriate approach. TS fuzzy models provide a powerful tool for modelling complex nonlinear systems because of its representational power, possibility of using learning algorithms to identify its parameters, gain insights into the local behaviour of the sub-models and its interpretability (Yen et al., 1998).

A variety of methods and architectures has emerged to determine the structure and parameters of TS fuzzy models. It is interesting to observe the progress in the field, particularly the evolution from a seminal knowledge-driven approach towards a data-driven approach. One of the main reasons for this is the permanent growth of available data from real-world applications and the increasing demands for effective decisions in real-time or almost real-time. These applications demand a new kind of admissible on-line learning algorithms in order to extract knowledge from data in a quick and efficient way.

The paper is organised in five sections. In section 2 the effluent neutralization process is described and some aspects related with its modelling are explained. Section 3 introduces the basic notions of the applied fuzzy modelling techniques. Experimental results for the second stage of the neutralization process are shown in section 4 and section 5 presents the conclusions.

2.   EFFLUENT NEUTRALIZATION PROCESS

Three different stages compose the mill effluent treatment plant: primary, secondary and tertiary stage. The primary treatment consists in the effluent neutralization and suspended solids removal, the secondary or biological treatment uses activated sludge for the degradation of dissolved organic matter and the tertiary treatment consists in colour removal and effluent equalization, regarding the quality and temperature.

In this work we focus on the primary treatment, more particularly on the study of the effluent neutralization process.

2.1 Description of the neutralization process

The acid effluent, the north alkaline effluent and the south alkaline effluent compose the plant effluent, Fig.1. Before the acid effluent enters the WWTP it is partially neutralised in two stages. The first stage consists in the addition of dust collected by the electro filter of the limekiln and the second stage consists in the addition of hydrated lime.
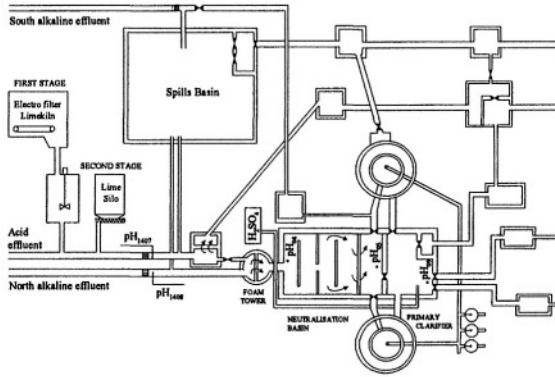

Figure 1. Plant of the effluent neutralization process.

The second stage has a much higher neutralization capacity than the first stage, though with associated cost, since hydrated lime has to be bought. The correct neutralization of the industrial effluent is vital for the functionality of the secondary treatment and the quality of the final effluent. The goal is to achieve at the end of the primary treatment a pH in the interval [6.0; 8.0] and simultaneously minimizing the consumption of neutralization materials.

2.2 Modelling of the neutralization process

2.2.1 Hierarchical structure

A hierarchical structure is proposed to deal with the high number of variables that influence the values of pH. Five sub-models are considered: first stage, second stage, foam tower, neutralization basin and primary clarifier. The hierarchical structure was created based on knowledge obtained from process engineers and operators. Only the most relevant variables have been considered. Fig. 2 represents the input and output variables for the different models.

2.2.2   Structure of the models

One of the goals of this work is to construct fuzzy models that are interpretable, simple and applicable. In order to achieve this goal, each sub-model will be a first-order model with time transport delay: the output y_k will depend only on one past value of each of the inputs, $u^p_{k-T_{d_p}}$, and one past value of the output, y_{k-1}, Eq. (1).

$$y_k = f\left(u^1_{k-T_{d_1}}, u^2_{k-T_{d_2}}, \ldots, u^p_{k-T_{d_p}}, y_{k-1}\right)$$ (1)
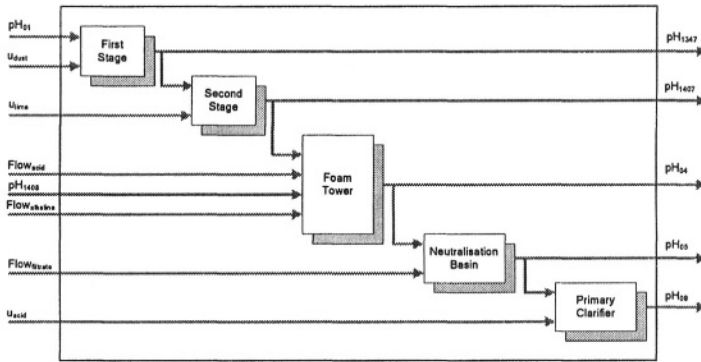

Figure 2. Hierarchical structure for the effluent neutralization process.

2.2.3   Input selection

The time transport delay for each variable was estimated based on process knowledge, completed with experimental data using ANFIS (Adaptive Network-based Fuzzy Inference System). ANFIS employs an efficient hybrid learning method that combines gradient descent and least-squares method to minimize the mean square modelling error (Jang, 1996).

For instance, in the second stage, the input variables are the pH of the acid effluent (pH_{1347}) and the addition of hydrated lime (u_{lime}). The output variable is the pH of the acid effluent (pH_{1407}), Fig. 2.

The dynamical model of the second stage can be expressed by:

$$y_k = f\left(u^1_{k-T_{d_1}}, u^2_{k-T_{d_2}}, y_{k-1}\right)$$ (2)

where the output y_k is a function of the two inputs, respectively at instants $k-T_{d_1}$ and $k-T_{d_1}$, and the output at instant k–1. To determine the time transport delay associated with the input variables $(pH_{1347}$ and $u_{lime})$ several fuzzy models were constructed, each with a different combination of past entries of these two inputs. For each input variable 10 candidate inputs $\left(u^1_{k-1}, u^1_{k-2}, \ldots, u^1_{k-10}; u^2_{k-1}, u^2_{k-2}, \ldots, u^2_{k-10}\right)$ were considered. The objective was to find the most influential two inputs, i.e., the ANFIS model with the smallest modelling error, using only a single epoch of training. Fig. 3 presents the values of the MSE (Mean Square Error) for each model. These values were calculated for the data set represented in Fig. 4.

The time transport delay for the two input variables is 2 minutes and 30 seconds, i.e., $T_d = 5$, since the sampling interval was $T = 30$ seconds.
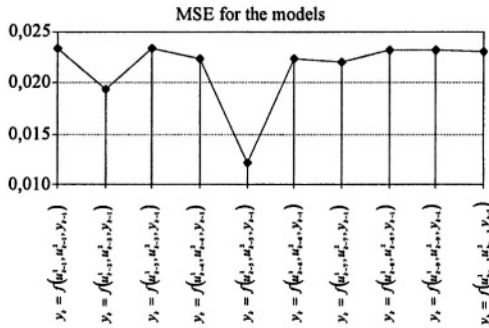

Figure 3. Time transport delay in the second stage of the process.

The dynamical model of the second stage can now be expressed by:

$$y_k = f\left(u^1_{k-5}, u^2_{k-5}, y_{k-1}\right) \tag{3}$$

The two input variables have the same time transport delay because it is impossible to dissociate the load of the process on the second stage, represented by the input variable pH1347, from the addition of hydrated lime.

3.   FUZZY MODELLING TECHNIQUES

This section briefly describes the two approaches used for modelling the effluent neutralization process: TS fuzzy models and eTS fuzzy models.

3.1   TS fuzzy models for off-line learning

TS fuzzy models were proposed in an effort to formalize a systematic approach to generate fuzzy rules from an input-output data set. The basic idea is to decompose the input space into fuzzy regions and to approximate the system in every region by a simple model (Takagi and Sugeno, 1985). A typical if-then rule in a first-order TS fuzzy model has the following format:

$$\Re_i : If \ x_1 \ is \ X_{i1} \ ...and \ x_n \ is \ X_{in} \ then \ y_i = a_{i0} + a_{i1}x_1 + \cdots + a_{in}x_n; \ i = 1,2,...,R \ (4)$$

where R_i denotes the ith fuzzy rule; R is the number of fuzzy rules; X_{ij} denotes the antecedent fuzzy sets, $j = 1,...,n$; y_i is the output of the ith linear subsystem and $\pi_i = [a_{i0}, a_{i1},..., a_{in}]^T$ are its consequent parameters.

The contribution of the corresponding local model to the overall output of the TS model is proportional to the firing strength τ_i of each rule, computed by Eq. (5), using algebraic product for fuzzy logic conjunction:

$$\tau_i = \mu_{i1}(x_1) \times \cdots \times \mu_{in}(x_n) = \prod_{j=1}^{n} \mu_{ij}(x_j); \ i = 1,2,...,R; \ j = 1,2,...,n \qquad (5)$$

Usually Gaussian membership functions are used to represent the antecedent fuzzy sets in such a way that a trade-off between complexity and precision of the model is obtained.

The TS model output is calculated by weighted averaging of the individual contribution of each rule:

$$y = \sum_{i=1}^{R} \lambda_i y_i = \sum_{i=1}^{R} \lambda_i x_e^T \pi_i \qquad (6)$$

where $\lambda_i = \dfrac{\tau_i}{\sum\limits_{i=1}^{R} \tau_i}$ is the normalized firing strength of the ith rule and x_e^T is

the extended input vector, $x_e^T = [1, x^T]$, allowing the free parameter in π.

The problem of identification of a TS fuzzy model is divided into two sub-tasks: learning the antecedent part of the model (consisting on the determination of the centres and spreads of the membership functions), and learning the parameters of the consequents. The identification on the antecedent part was solved by subtractive clustering the input-output space. This approach can be used for initial estimation of the antecedent parameters

and the resulting cluster centres are used as parameters of the antecedent parts, defining the centres of the rules of the TS fuzzy model (Chiu, 1994). The estimation of the parameters of the consequents is a least squares problem for fixed antecedent parameters. The vector of linear sub-model is made locally optimal in least squares sense (Angelov and Filev, 2004).

3.2   eTS fuzzy models for on-line learning

eTS fuzzy models have been recently introduced by Angelov and Filev and the approach is based on an on-line learning algorithm that recursively develops the model structure and parameters (Angelov and Filev, 2004). Structure identification includes estimation of the antecedent parameters by recursive fuzzy clustering and the consequent parameters are obtained by applying the modified Recursive Least Squares (w)RLS estimation algorithm. In the following the basics of the algorithm are briefly recalled.

3.2.1   On-line estimation of the antecedent parameters

The recursive fuzzy clustering algorithm uses the notion of the informative potential (accumulated spatial proximity measure), introduced in the mountain clustering algorithm and then refined in the subtractive clustering algorithm, to compute the potential of each new data point in order to select the cluster centres (focals).

The algorithm starts with the first data point established as the focal point of the first cluster. Its coordinates are used to form the antecedent part of the fuzzy rule using Gaussian membership functions and its potential is assumed equal to 1. Starting from the next data point onwards the potential of the new data points is calculated recursively. In contrast with the subtractive clustering algorithm there is not a specific amount subtracted from the highest potential, but update of all the potentials after a new data point is available on-line. The potential of the new data sample z_k is recursively calculated as Eq. (7) (Angelov and Filev, 2004),

$$P_k(z_k) = \frac{k-1}{(k-1)(\vartheta_k + 1) + \sigma_k - 2v_k} \tag{7}$$

where $\vartheta_k = \sum_{j=1}^{n+1}(z_k^j)^2$, $\sigma_k = \sum_{l=1}^{k-1}\sum_{j=1}^{n+1}(z_l^j)^2$, $v_k = \sum_{j=1}^{n+1}z_k^j\beta_k^j$ and $\beta_k^j = \sum_{l=1}^{k-1}z_l^j$.

Parameters ϑ_k and v_k are calculated from the current data point, z_k, while β_k^j and σ_k are recursively updated by Eq. (8) (Angelov and Filev, 2004).

$$\sigma_k = \sigma_{k-1} + \sum_{j=1}^{n+1} \left(z_{k-1}^j\right)^2 ; \; \beta_k^j = \beta_{k-1}^j + z_{k-1}^j \tag{8}$$

After the new data are available they influence also the potentials of the centres of the clusters $(z_l^*, l = 1,\ldots,R)$, which are respective to the focal points of the existing clusters $(x_l^*, l = 1,\ldots,R)$. The reason is that by definition the potential depends on the distance to all data points, including the new ones. The potential of the focal points of the existing clusters is recursively updated by Eq. (9) (Angelov and Filev, 2004).

$$P_k\left(z_l^*\right) = \frac{(k-1)P_{k-1}\left(z_l^*\right)}{k - 2 + P_{k-1}\left(z_l^*\right) + P_{k-1}\left(z_l^*\right)\sum_{j=1}^{n+1}\left(d_{k(k-1)}^j\right)^2} \tag{9}$$

where $P_k(z_l^*)$ is the potential of the cluster z_l^* at time k, which is a prototype of the lth rule, $d_{k(k-1)}^j = z_k^j - z_{k-1}^j$, denotes the projection of the distance between two data points, z_k^j and z_{k-1}^j, on the axis z^j (x^j for $j = 1,2,\ldots,n$ and on the axis y for $j = n + 1$).

Further developing the previous algorithm, if the potential of the new data point is higher than the maximum potential of the existing centres or lower than the minimum then the new data point is accepted as a new centre as Eq. (10):

$$\text{If } P_k(z_k) > \overline{P}_k \text{ or } P_k(z_k) < \underline{P}_k \text{ then } R = R+1; x_R^* = x_k; P_k(x_R^*) = P_k(x_k) \tag{10}$$

where $\overline{P}_k = \max_{l=1}^{R} P_k(z_l^*)$ and $\underline{P}_k = \min_{l=1}^{R} P_k(z_l^*)$.

If in addition to the previous condition the new data point is close to an old centre then the new data point replaces this centre, instead of being accepted as a new one, Eq. (11):

$$\text{If } \left(P_k(z_k) > \overline{P}_k \text{ or } P_k(z_k) < \underline{P}_k\right) \text{ and } \delta_{min} < \frac{r}{2} \text{ then}$$
$$z_j^* = z_k; P_k(z_j^*) = P_k(z_k) \tag{11}$$

where $\delta_{min} = \min\limits_{l=1}^{R} \left\| z_k - z_l^* \right\|$, r is the constant radii and j is the index of the replaced centre (closest cluster centre to the new data point).

3.2.2 On-line estimation of the consequent parameters

The problem of increasing size of the training data is handled by the RLS algorithm for the globally optimal case and by the weighted RLS algorithm for the locally optimal case. Under the assumption of a constant/unchanged rule base the optimization problems are linear in parameters.

In eTS fuzzy models, however, the rule base is assumed to be gradually evolving. Therefore, the number of rules as well as the parameters of the antecedent part will vary. Because of this evolution, the normalized firing strengths of the rules will change. Since this affects all the data (including the data collected before time of the change) the straightforward application of the RLS or wRLS algorithm is not correct. A proper resetting of the covariance matrices and parameters initialization of the algorithms is needed at each time a rule is added to and/or removed from the rule base. The modified RLS algorithm for global and local parameter estimation is described in the reference (Angelov and Filev, 2004).

3.2.3 Rule base evolution in eTS fuzzy models

The recursive procedure for on-line learning of eTS fuzzy models includes the following stages:
1. Initialization of the rule base structure (antecedent part of the rules);
2. Reading the next data sample at the next time step;
3. Recursive calculation of the potential of each new data sample, Eq. (7);
4. Recursive update of the potentials of old centres, Eq. (9);
5. Possible modification or upgrade of the rule base structure, Eqs. (10)-(11);
6. Recursive calculation of the consequent parameters;
7. Prediction of the model output for the next time step.
   The algorithm continues from stage 2 at the next time step.

4. EXPERIMENTS

This section describes the experiments done in the second stage of the effluent neutralization process to predict the pH values.

4.1   Experiment 1

In the first experiment, two plant data sets from consecutive days are used, the first one, Fig. 4, for training, and the second one, Fig. 5, to determine the generalisation capability of the fuzzy models (validation). Each data set has 2880 samples and the data is normalized into de domain [0, 1] to avoid negative influence in clustering results from variations in the numerical ranges of the different features (Babuska, 1998).
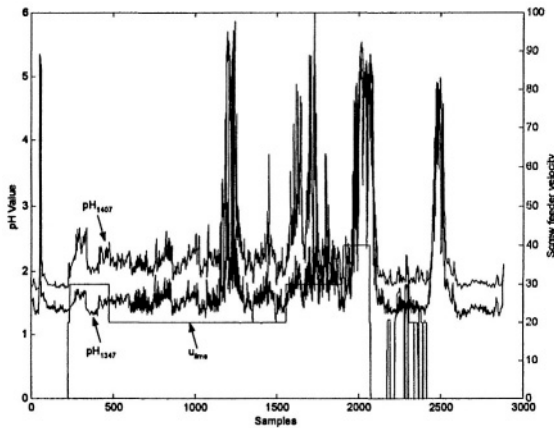

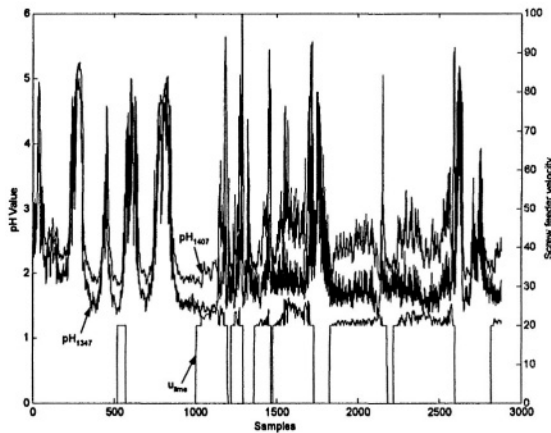*Figure 4.* Training data set for the second stage of the effluent neutralization process.

Figure 5. Validation data set for the second stage of the effluent neutralization process.

4.1.1 TS fuzzy model

The TS fuzzy model approach is applied for off-line learning. The value for the constant radii was, $r = 0.3$.

Fig. 6 shows the evolution of the process and model output for training (left) and validation (right). Table 1 presents the parameters of the TS fuzzy model acquired from data, namely information about the centres of the Gaussian membership functions and the parameters of the linear sub-models. The width of the membership functions is the same for all the fuzzy sets, $\sigma_{ij} = 0.1061$.
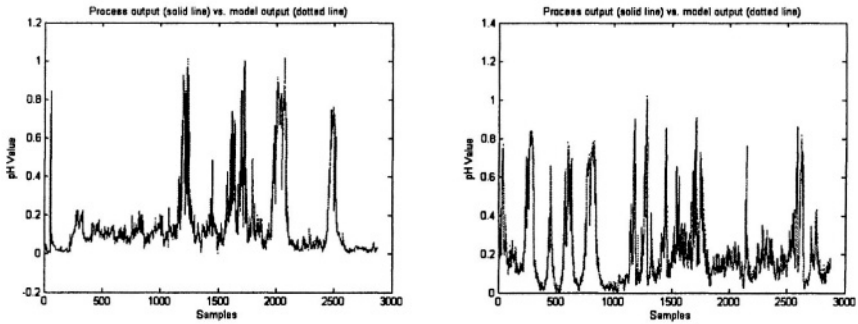

Figure 6. Off-line learning with TS fuzzy model approach.

Even with a small number of fuzzy rules, only 3, the performance of the TS fuzzy model is satisfactory both for training and validation data. It must be stated that the structure and parameters of the TS fuzzy model obtained from training data remain unchanged when the model is tested with the validation data. The accuracy of the model can be improved if a smaller value for radii is considered since this parameter can be used to control the number of fuzzy rules and inherently the model complexity and precision.

Table 1. TS fuzzy model parameters, off-line learning.

Rule	u^1_{k-5}	u^2_{k-5}	y_{k-1}	y_k
R_1	0.0746	0.4989	0.1045	$y^1 = 0.7207 - 0.4268u^1_{k-5} + 0.6361u^2_{k-5} + 0.1909y_{k-1}$
R_2	0.0565	0	0.0296	$y^2 = 0.4170 - 0.1545u^1_{k-5} + 0.5513lu^2_{k-5} - 0.0108y_{k-1}$
R_3	0.0847	0.7496	0.1348	$y^3 = 0.2721 - 0.1682u^1_{k-5} + 0.75261u^2_{k-5} + 0.1443y_{k-1}$

4.1.2 eTS fuzzy model

The eTS fuzzy model approach is applied for on-line learning. The value for the constant radii was $r = 0.5$ and the initialization parameter for RLS

was $\Omega = 750$. Fig. 7 shows the evolution of the process and model output for training (left) and validation (right). Table 2 presents the parameters of the eTS fuzzy model at the end of the learning process for training data.

The structure and parameters of the eTS fuzzy model, at the end of the training, were exactly the same with which the eTS fuzzy model was initialized for validation. The number of rules at the end of training was 8, and it remains the same after validation, but some fuzzy rules were replaced and the parameters of the model were constantly updated. In practice what happens is that at every instant a new model is constructed and even though the structure (number of fuzzy rules) does not change the model parameters are varying.

The performance of the eTS fuzzy model is satisfactory but the complexity of the eTS fuzzy model is higher than the TS fuzzy model. Another important aspect is the poor interpretability of the model since there are a lot of similar membership functions, Table 2.
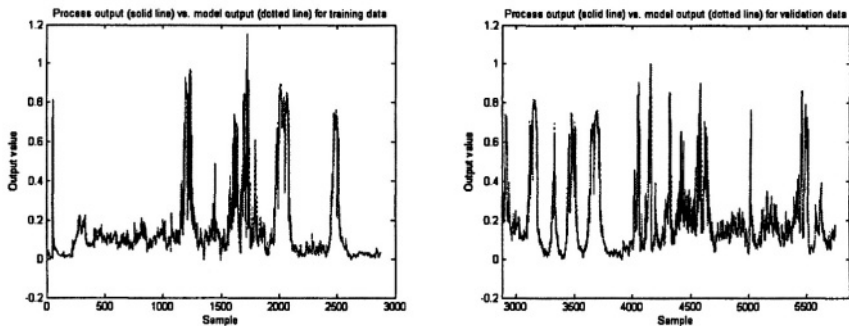

Figure 7. On-line learning with eTS fuzzy model approach.

Table 2. eTS fuzzy model parameters, on-line learning.

Rule	u_{k-5}^1	u_{k-5}^2	y_{k-1}	y_k
R_1	0.0795	0	0.0264	$y^1 = -0.0067 + 0.4768u_{k-5}^1 + 0.0183u_{k-5}^2 + 0.3773y_{k-1}$
R_2	0.5232	0	0.0433	$y^2 = -0.0471 + 0.4566u_{k-5}^1 + 1.3214u_{k-5}^2 + 0.5224y_{k-1}$
R_3	0.8962	0	0.1145	$y^3 = 0.6906 - 0.6354u_{k-5}^1 - 0.6062u_{k-5}^2 + 1.4509y_{k-1}$
R_4	0.9545	0	0.3133	$y^4 = -1.1746 + 1.4509u_{k-5}^1 + 0.7195u_{k-5}^2 + 1.4558y_{k-1}$
R_5	0.9809	0	0.8124	$y^5 = -0.2341 + 0.0694u_{k-5}^1 + 0.0401u_{k-5}^2 + 1.2025y_{k-1}$
R_6	0.0635	0.7486	0.1169	$y^6 = 0.0899 + 0.1542u_{k-5}^1 - 0.1045u_{k-5}^2 + 0.8327y_{k-1}$
R_7	0.0922	0.4961	0.1261	$y^7 = -0.0107 + 1.0377u_{k-5}^1 - 0.0803u_{k-5}^2 + 0.5816y_{k-1}$
R_8	0.9933	0.9956	0.8806	$y^8 = 0.1300 + 0.3535u_{k-5}^1 + 0.0630u_{k-5}^2 + 0.3808y_{k-1}$

4.2   Experiment 2

In order to more effectively test the generalisation capabilities of the fuzzy models a second validation data set, from a third day, was considered, Fig. 8. This data set is from a different month and it can be observed that the input variables present a quite different behaviour.
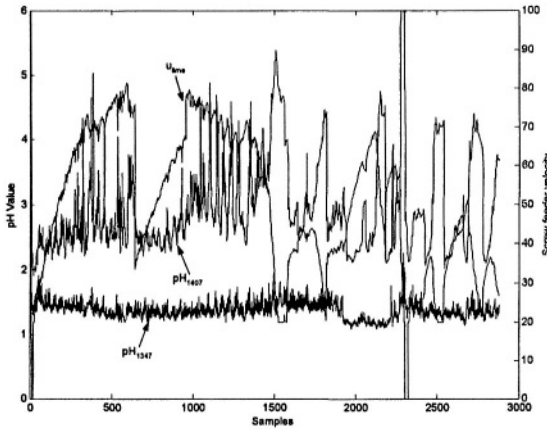

Figure 8. Validation data set (2) for the second stage of the effluent neutralization process.

4.2.1   TS fuzzy model

The behaviour of the TS fuzzy model is presented in Fig. 9 and it is clear that the error between the process and model output has increased considerably. This is because the dynamics of the process has changed and the structure and parameters of the model remained fixed (the same as in experiment 1).
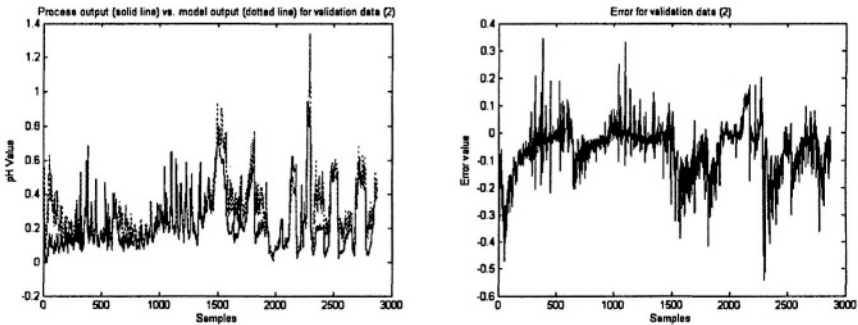

Figure 9. Off-line learning with TS fuzzy model approach for validation data set (2).

This confirms what we already know, process dynamics change with time and for off-line learning the only way to improve on this is to use more representative data or models must be build for different operating points.

4.2.2   eTS fuzzy model

Fig. 10 presents the behaviour of the eTS fuzzy model and it is satisfactory, particularly for the second half of the data set. The number of fuzzy rules at the end is the same as in experiment 1 but the number of replaced fuzzy rules has increased, with some rules being replaced more than once.
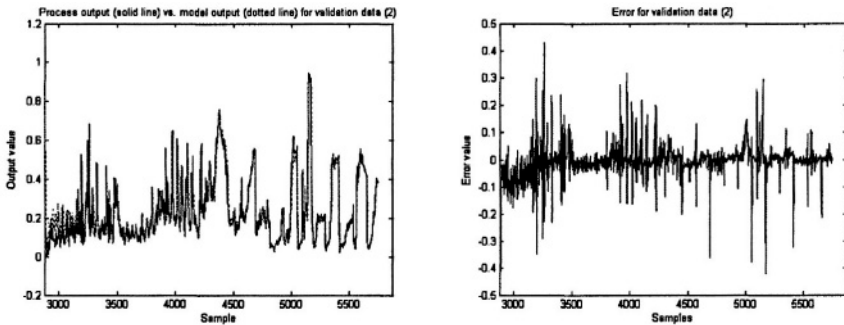

Figure 10. On-line learning with eTS fuzzy model approach for validation data set(2).

With on-line learning even if the training data is not representative of all the operating points of the process it is possible to achieve satisfactory results since the approach can evolve the structure and parameters of the model, according with the current dynamics of the process. The interpretability of the model remains poor and to improve on this rule base simplification and reduction methods must be applied during the learning process (Victor Ramos and Dourado, 2004; Setnes et al., 1988).

5.   CONCLUSIONS

In this paper the modelling of the effluent neutralization process of a WasteWater Treatment Plant in pulp industry is presented. Fuzzy modelling techniques for off-line learning and on-line learning were used and on-line learning of fuzzy rules proves to be an efficient technique for building up predictive models, particularly when processes present non-linear characteristics and time-varying dynamics. These models will be used for monitoring and for synthesizing automatic closed-loop controllers.

Although some improvements on the interpretability of the models are necessary their adaptive nature makes them a useful tool for on-line modelling of real-world applications of AI in decision, monitoring and control.

ACKNOWLEDGEMENTS

REFERENCES

Angelov, P.; Filev, D., 2004, An Approach to Online Identification of Takagi-Sugeno Fuzzy Models, *IEEE Transactions on Systems, Man, and Cybernetics.* **34**(1): 484-498.

Babuska, R., 1988, Fuzzy Modeling for Control, *International Series in Intelligent Technologies,* Kluwer Academic Publishers.

Chiu, S., 1994, Fuzzy Model Identification Based on Clustering Estimation, *Journal of Intelligent and Fuzzy Systems,* **2**: 267-278.

Jang, J.-S. R, 1996, Input Selection for ANFIS Learning. in: *Proceedings of the IEEE International Conference on Fuzzy Systems,* New Orleans.

Setnes, M., 2001, Complexity Reduction in Fuzzy Systems, *PhD Thesis,* Delft University of Technology.

Setnes, M, Babuska, R., Kaymak, U., and van Nauta Lemke, H.R., 1988, Similarity Measures in Fuzzy Rule Base Simplification, *IEEE Transactions on Systems, Man, and Cybernetics,* **28**(3): 376-386.

Victor Ramos, J., and Dourado, A., 2004, On-Line Interpretability by Fuzzy Rule Base Simplification and Reduction, *EUNITE Conference,* Aachen, Germany.

Takagi, T., and Sugeno, M., 1985, Fuzzy Identification of Systems and Its Applications to Modeling and Control, *IEEE Transactions on Systems, Man, and Cybernetics,* **15**(1):116-132.

Yen, J., Wang, L., and Gillespie, W., 1998, Improving the Interpretability of TSK Fuzzy Models by Combining Global and Local Learning, *IEEE Transactions on Fuzzy Systems,* **6**(4):530-537.