Intelligent Control at the Coagulation Process in a Drinking Water Treatment Plant

Hector Hernandez[1], Madaín Pérez[1], Jorge Camas[1], Rafael Mota[1], Nicolás Juárez[1] and Marie-Véronique Le Lann[2]

[1]Instituto Tecnológico de Tuxtla, ITTG
Carretera Panam. Km. 1080, 29050-Tuxtla Gutiérrez, Chiapas, México
(hhernandezd,mperez,jcamas,rmota,njuarez)@ittg.edu.mx
http://www.ittg.edu.mx

[2]INSA, Département de Génie Electrique et Informatique
135, Avenue de Rangueil 31077 Toulouse cedex 4 France
mvlelann@laas.fr
http://www.insa-tlse.fr

Abstract. With increasing demands for high precision autonomous control over wide operating envelopes, conventional control engineering approaches are unable to adequately deal with system complexity, nonlinearities, and temporal parameter variations, and with uncertainty. Intelligent Control or self-organising/learning control is a new emerging discipline that is designed to deal with problems. The coagulation unit is a major step in the production of potable water, allowing the removal of colloidal particles and contamination sources. In order to obtain a simple model to describe the water treatment plant, a behavior model sets out, from the analysis of raw water characteristics to the entrance of the plant: (1) to develop a software sensor based on artificial neural networks for predicting on-line the amount of optimal coagulant dosage, and (2) the determination of the functional states in real time (diagnosis system based on fuzzy classifier).

1 Introduction

The water industry is facing increased pressure to produce higher quality treated water at a lower cost. The drinking water plant object of this study is the "SMAPA" (SMAPA, 2007[1]) of Tuxtla city in Mexico. The control of the plant is fundamental to maintain a good quality of service. This has motivated important efforts in the development of the methods of control and automatic monitoring in the last few years [2]. Coagulation process is one of the most important stages in surface water treatment, allowing the removal of colloidal particles [3].

This paper addresses the problem of coagulation control based on the raw water characteristics such as turbidity, temperature and pH, in a global system including the analysis and the determination of the functional states and the detection of fault. Coagulant dosing is not only one of the major control parameter in coagulation

© G. Sidorov, M. Aldape, M. Martinez, S. Torres. (Eds.) *Received 09/03/09*
Advances in Computer Science and Engineering. *Accepted 27/04/09*
Research in Computing Science 42, 2009, pp. 347-358 *Final version 06/05/09*

process, but also the major operation cost in water treatment plant. Most coagulant dosing is determined by the way of jar test. However, the jar test can only provided periodic operation information, which can not be applied to real-time control of the coagulation process, especially with a time-varying raw water quality (principally during a spring-summer runoff period of the time). Excessive coagulant overdosing leads to increased treatment costs and public health concerns, while underdosing leads to a failure to meet the water quality targets and less efficient operation of the water treatment plant.

One of the fields of machine learning more developed is the classification. This method, based on Fuzzy Logic, presents many advantages which make it well adapted to chemical or biotechnological complex processes. In the monitoring of dynamic processes, the state of operation of the system determines the situation or operating condition. The objects to be classified are the situations observed in real time. Good coagulation control is essential for water quality and economic plant operation [4], [5]. Finding a general mathematical model for biotechnological process is somehow complex. During the last decade a number de models based on artificial neuronal networks (ANN) have been developed and applied for predicting coagulant dose concentrations of water treatment process (Baba, 1990 [6]; Mirepassi, Cathers and Dhamarppa, 1995 [7]; A fuzzy neural system, coupling of ANN with fuzzy theory, has been applied to extract the control rules by learning story operational data of water treatment process [8]. Another works based on ANN's involved in the production of potable water have been carried out (Fletcher et al. 2001[9]; Baxter et al., 2002[10]; Peijin and Cox, 2004[11]).

This paper addresses the problem of automatic coagulation control based on the raw water characteristics such as turbidity, pH, temperature, etc. Some recent studies (Valentin, PhD thesis 2000 [3]; Lamrini and LeLann, 2004 [2]; and Hernandez, PhD thesis 2006 [12]) have shown the potential effectiveness of such an approach based on ANN's by means of the implementation of a neural software sensor for on-line prediction of the coagulant dosage. The innovative aspect of this work resides in the integration of various techniques in a global system including data pre-processing, automatic control of coagulation, analysis of uncertainties and the possibility of integration as entry to a system of diagnosis, which should allow the portability of the system at low cost from one site to another.

A brief description of the water treatment plant and the operation units involved in the drinking water treatment process is first provided in section 2. The LAMDA technique of classification is described in section 3. The methodology used for the design and synthesis of the Neural Software Sensor and system diagnosis is given in section 4. Finally, experimental results are presented and discussed in section 5.

2 Water Treatment Process

2.1 Overview of Water Treatment Operations

The water is the most abundant compound on the surface of the world. Water treatment involves physical, chemical and biological processes that transform raw water into drinking water which satisfies a whole of standards of quality at a reasonable price for the consumer.

The "SMAPA" water treatment plant (Tuxtla city, Mexico), which was used as an application site for this study, provides water to more than 800,000 inhabitants and has a nominal capacity to process 800 l/s of water per day. The figure 1 presents a schematic overview of the various operations necessary to treat the water, the available measurements, and the coagulant dosing point. The complete usual chain comprises the 5 great following units: pre-treatment, pre-oxidation, clarification, disinfection, and refining. The present work concerns essentially the coagulation process. Raw water is abstracted from the river "Grijalva" and pumped to the treatment works. Water treatment plants invariably include two main process units, clarification and filtration. Other units may be required depending of the quality of the water source. The coagulation process is brought about by adding a highly ionic salt (aluminum sulphate) to the water.

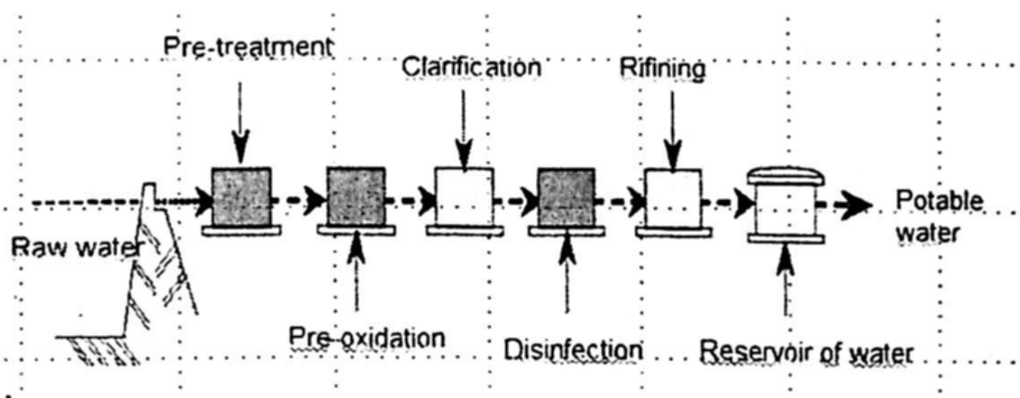

Figure 1. Potable water plant.

A bulky precipitate is formed which electrochemically attracts solids and colloidal particles. The solid precipitate is removed by allowing it to settle to the bottom of the tank and then periodically removing it as sludge. The coagulation process accounts for the removal of most of the undesirable substances from the raw water and hence tight monitoring and control of this process is essential. The next stage is filtration, where the particles passing trough the previous stages are removed. The final stages in the process are chlorination and pH adjustment. The water is then stored in a tank and ready to be transported through the water supply network.

2.2 Coagulation Control

Surface waters contain both dissolved and suspended particles. The suspended particles vary considerably in source, composition charge, particle size, shape and density. The correct design of a coagulation process and the selection of coagulants depend upon understanding the interactions between these factors. This process is one of the most important stages in surface water treatment, allowing for the removal of colloidal particles. The main difficulty is to determine the optimum quantity of chemical reagent related to raw water characteristics. Poor control leads to wastage of expensive chemicals, failure to meet the water quality targets, and reduced efficiency of sedimentation and filtration processes. In contrast, good control can reduce manpower and chemical costs and improve compliance with treated water quality targets. The traditional method of controlling coagulant dose, called the jar-test, relies heavily upon human intervention. It involves taking raw water samples and applying different quantities of coagulant to each sample. After a short period of time each sample is assessed for water quality and the dosage that produces the optimal result is used as a set point. Operators change the dose and make a new jar test if the quality of treated water changes. Disadvantages associated with such a procedure are the necessity to rely on manual intervention, and lack of adaptation to abrupt changes of water characteristics. The objective of this paper is to provide a complementary support to the jar-test allowing for the automatic determination of optimal coagulant dose from raw water characteristics, using an artificial network approach. This approach requires the availability of on-line water quality measurements at an upstream survey station.

3 LAMDA classification technique

LAMDA (Learning Algorithm for Multivariate Data Analysis) methodology is a classification technique developed by Joseph Aguilar-Martin and others in LAAS-CNRS [13]. Previous works [14], [15], [16], [17], described in detail the methodology, as well as the algorithms and functions used. We will limit ourselves in this work to present the main characteristics of the methodology and its general operation. LAMDA algorithm represents a system of classes or concepts by means of the logic connection of all marginal information available [18]. LAMDA is a fuzzy methodology of conceptual clustering and classification, in this way, the global adequacy of an object (individual) to a class is equivalent to its membership to the class, and it is calculated from the marginal adequacy of each attribute and according to the heuristic rule of Maximum Adequacy. An object is then assigned to the class which presents the greater adequacy degree. The total indistinguishability or homogeneity inside the description space from which the information is extracted is introduced by means of a special class called the Non-Informative Class (NIC), this class accepts the same adequacy degree any object of the description space, and a minimum classification threshold is therefore induced.

Let us consider a collection of objects or situations X, and a finite set of n qualitative or quantitative descriptors, **A**. An object is represented by a vector $\underline{x}$ and its j^{th} component is the value taken by the j^{th} descriptor. The information conveyed by each descriptor contributes to the membership of the element to the class by means of the Marginal Adequacy Degree (MAD). The Global Adequacy Degree (GAD) to a class is a Fuzzy Logic combination of the MAD's as shown in figure 2. Piera and Aguilar-Martin (1991) introduced the mixed connectives of lineal compensation that interpolate between a conjunctive and a disjunctive logic operator. It was shown that such interpolation is completely ordered with respect to the "exigency degree", the highest in the conjunction case (AND) and the lowest in the disjunction one.

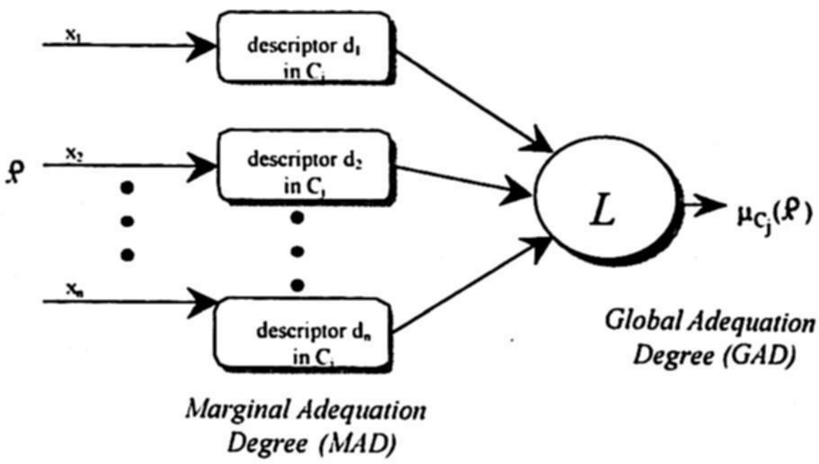

Figure 2. The Marginal and Global Adequacy Degree (MAD and GAD).

The computation for the class parameters is done independently for each component, using separately the measurement given by each sensor. Only averaging functions are needed for that algorithmic step. Each vector x, or situation is assigned to one of the existing classes according that its adequacy GAD exceeds the NIC adequacy, otherwise this object is assigned to the NIC class in case of passive recognition. In the learning case, this vector starts the creation of a new class and contributes to its initialisation; therefore the previous knowledge of the number of classes in not needed. Waissman et al. (1998) have been proposed several MAD functions, as well as their corresponding sequential learning algorithms.

4 Methodology of Neural Software Sensor and the System Diagnosis

The coagulation process is difficult to model using traditional models. The coagulant dosage ensuring optimal treatment efficiency has been shown experimentally to be non-linearly correlated to raw water characteristics which are usually available on-line. The system developed for the prediction of the optimal coagulant dosage and the system diagnosis it was divided into two modules: (1) Determination of coagulant dosage using Artificial Neural Networks (software sensor), (2) System diagnosis that allows the process expert to obtain the classification that represents the best the process situations in order to use it later for pattern recognition.

4.1 Mode of the Prediction of Optimal Coagulation Dosage

The general method for the prediction of coagulant dose is shown in Figure 3. We analyze the following stages: pretreatment of data descriptors of the water by using ACP (principal component analysis), and the implementation of an iterative procedure to test the data for training and testing determining a confidence interval of prediction using the bootstrap technique (software sensor).

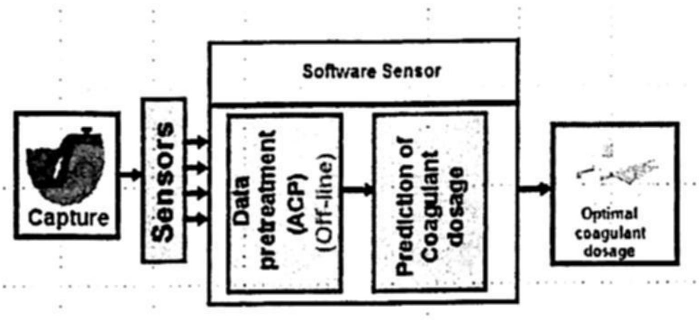

Figure 3. Model of the prediction of optimal coagulation dosage.

4.2 Principal Components Analysis (PCA)

PCA is one of the multivariate methods of analysis and has been used widely with large multidimensional data sets. The PCA method is applied to determine the main characteristic of the variables necessary for the prediction of the optimal coagulant dosage. These characteristic variables are considered as the input variables to the neural model for which the training algorithm is performed. The PCA generates a new set of variables called principal components. Each principal component is a linear combination of the original variables. All the principal components are orthogonal to each other so there is not redundant information. The principal components as a whole form an orthogonal basis for the data space. The first principal component is a single axis in space. When each observation is projected on that axis, the resulting values form a new variable, and the variance of this variable is the maximum among all possible choices of the first axis. The second principal component is another axis in space, perpendicular to the first. Projecting the observations on this axis generates another new variable. The variance of this variable is the maximum among all possible choices of this second axis. The full set of principal components is as large as the original set of variables. But is common for the sum of the variances of the first few principal components to exceed 80% of the total variance of original data.

4.3 Artificial Neural Networks (ANN's) and Multi-Layer Perceptron (MLP)

ANN's are one of the earliest adaptive techniques in engineering and computing science. The concept of ANN was inspired by the way of the biological brain processes information. An ANN is a network of neurons or processing elements and weighted connections [19]. ANN is fundamentally a mathematical model composed by a set of nodes (artificial neurons) where information is processed. An ANN can be classified into two different categories: unsupervised or supervised learning. In the situation of an unsupervised model, the networks seek to identify features of the training patterns without external assistance. On the other hand, for the supervised learning process is necessary to use in the train input-output patterns. For each input pattern, the network generates an output pattern, which is compared with the desired output and the adaptation of the model parameters is made in relation to the observed error. One of the most studied and used ANN architecture is the Multi-Layer Perceptron (MLP). The prediction of optimal coagulant dosage from water characteristics is a non linear regression problem which can be tackled using MLP's. Consists of an input-output network, which have the neurons distributed by several layers, fully connected between adjacent layers, and where the flow of information is done in a feed-forward way. The MLP is usually trained by gradient descent methods [20], in which the error is propagated backwards through the network. Figure 4 shows a MLP with three layers: an input layer (variables of the raw water quality parameters) with n neurons, a hidden layer with H neurons and a layer with one output neuron (variable of the optimal coagulant dosing rate).

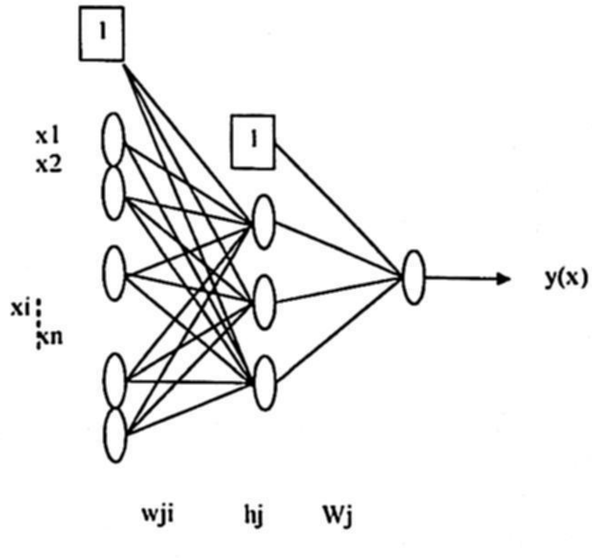

Figure 4. Multi-Layer Perceptron (MLP).

The expression of the output of MLP is given by:

$$y(x) = \sum_{j=1}^{H} W_j h_j + W_0 \quad \text{and} \quad h_j = \sum_{i=1}^{n} w_{ji} x_i + w_{j0}$$

Where: W_{ji} are the weights between the input layer and the hidden layer and W_j the weights between the hidden layers and the output layer.

The input nodes do not make any kind of processing and sending the input patterns to the first hidden layer is their only function. Conversely, the neurons from the other layers have the capacity of processing the received information. Each one of them performs two different operations: the weighted sum of its inputs (using the weights associated with the existing links between this neuron and the others from the previous layer), followed by a non-linear transformation (called by activation function or transfer function). The resulting output from these two actions is then sent on to the next layer. To sum up, if we have a MLP such as the one represented in figure 2 and with the same activation function, h , in all its neurons, then it can be described mathematically as:

$$h_j = \sum_{i=1}^{n} w_{ji} x_i + w_{j0}$$

The transfer function can be any function, but for most practical uses of neural networks it is important to have a continuous, completely differentiable function. Over the years many transfer functions have been proposed [21], but the most prominent ones for neural networks are linear and transfer function. In our case, we know that the relationship between the coagulant dosage with the raw water characteristics is non-linear so the choice of a sigmoid transfer function has been made. The method traditionally used to perform the training of such networks, e.g. to adjust the weighted connections, is the Backpropagation learning algorithm. The term Backpropagation refers generally to the manner in which the gradient is computed for non-linear Multi-layer networks. There are a number of variations on the basic algorithm which are based on other standard optimization techniques, such as conjugate gradient, Newton and Levenberg-Marquardt methods. Learning occurs when the network emulates the non-linear function underlying the training data set. The weighted connections are adjusted by minimizing the following criteria derived from the difference between real and neural outputs respectively t and y, as:

$$C = \frac{1}{N} \sum_{i=1}^{N} (t_i - y_i)^2$$

4.4 System Diagnosis

This part is a general procedure for monitoring the plant. It comprises the following modules: (1) Data pretreatment, (2) The method of form recognition (fuzzy classifier) to generate the class association to functional states of the process (Figure 5) and considering the stage of recognition in real time (Figure 6). At this part it is very important the active participation of the expert, both in the learning phase and the online recognition phase, which displays the current status of the water treatment plant.

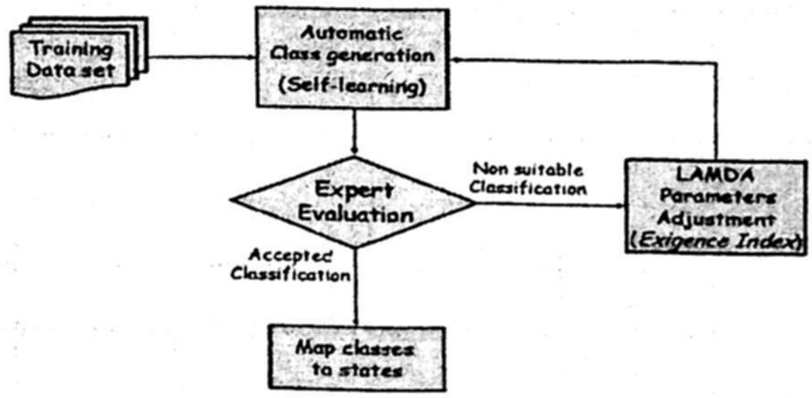

Figure 5. Association classes to functional states (analysis of historical data-non-supervised learning).

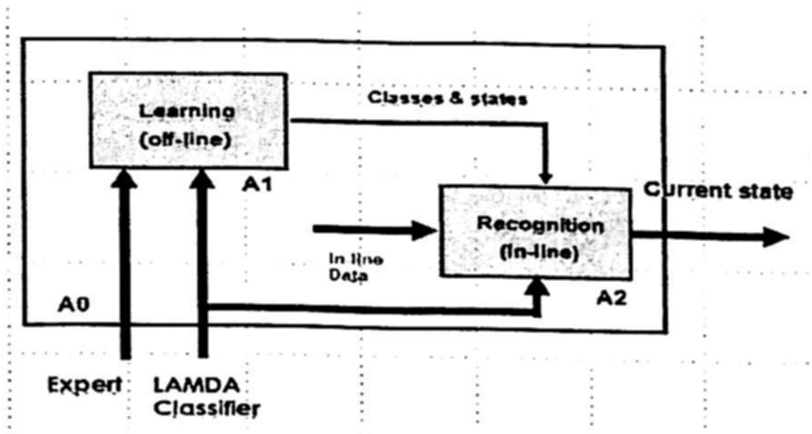

Figure 6. In-line functional states recognition (LAMDA fuzzy classifier).

5 Application the Method to the SMAPA Plant

The raw database consists of 728 measurements of 9 variables during a period of 24 months (2006-2007). Every sample underwent to different physical and chemical analysis as well as to the jar-testing in order to determine the coagulant dosage.

5.1 Prediction of Coagulant Dosage

After implementing the ACP, the number of input neurons is 4 (TUR, TC, TEMP and pH). The training was carried out over the first year (2006). The average error represented by the Matlab criterion *MSE* [22] calculated by the network is 0.085 on the training set corresponding to data of year 2007. The validation of the ANN has been performed on test data of year 2007 not included in the initial training set: the criterion *MSE* is a little higher (0.092). Figure 7 shows the prediction accuracy of the ANN model for this validation set of year 2007. The predictions given by neural network (point line) are very close to the real data.

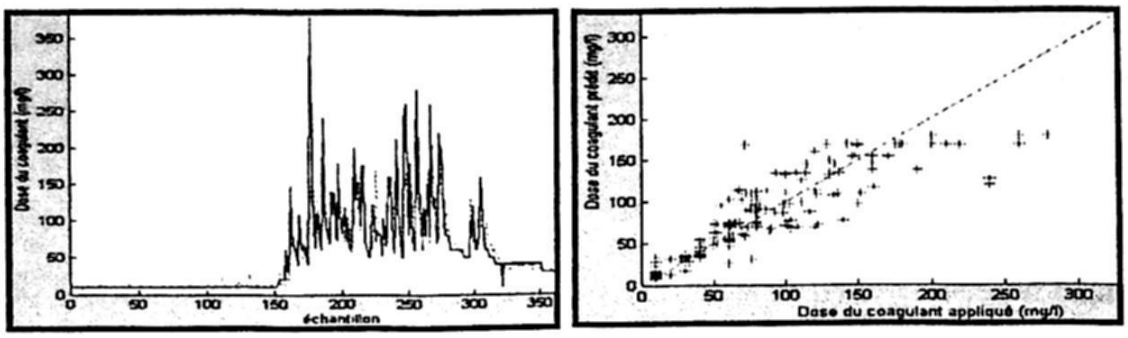

Figure 7. Left: actual (thin line) versus predicted (point line) coagulant dosage with ANN model on test data. Right: predicted versus target coagulant dose.

ANN model is compared with a multi-linear decomposition of the dose versus with the same input variables. Prediction accuracy is clearly poorer than the one of the neuronal model performance.

5.2 Results with the system diagnosis

Functional states of the water treatment process and according to the method proposed in the previous section, using the LAMDA classification algorithm as a general strategy, for obtaining the model of the plant (using historical data), as in recognition of the functional states in real time according to online measurement of the variable characteristics of the water at the entrance to the plant. Figure 8 shows the results for 5 classes. These classes are associated with 5 functional states (shown in figure) with the assistance of the expert from the drinking water treatment.

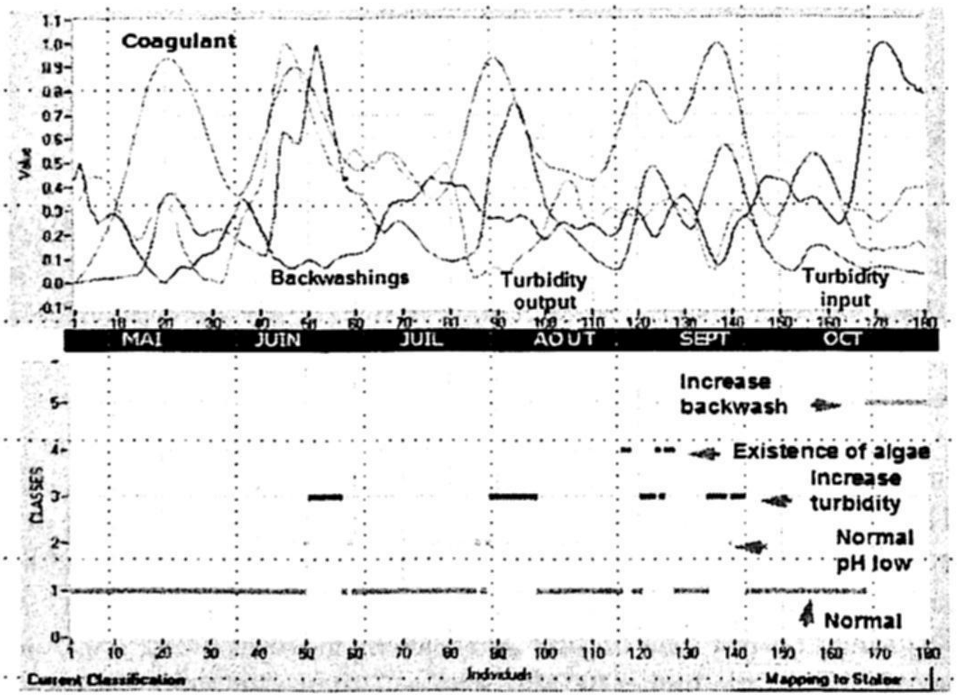

Figure 8. In-line functional states recognition (LAMDA fuzzy classifier) [23].

6 Conclusions

A neural software sensor for coagulation control was developed from Artificial Neural Networks. He supplies us on real time the coagulant dosage to inject in the coagulation unit, key stage of the process in a water treatment plant. The selection of the entries of the network was made by using the statistical technique ACP to allow eliminating the redundant information. Experimental results using historical real data have demonstrated the efficiency of this approach.

The LAMDA fuzzy classifier technique has been used to help to the interpretation of operational behaviour of a potable water treatment plant (SMAPA plant unit of Tuxtla Gutierrez, Mexico). Their classification under the unsupervised mode (i.e. with automatic class generation but controlled by the expert through the parameter α) exhibited 5 classes which could be aggregated to 5 functional states. This paper showed the realization of an intelligent monitoring system, where the reasoning task consists in the combination of data inspection and classification situations and of the expert knowledge. This strategy will be applied to the monitoring of the overall water treatment plant with measured data from different units in the plant (not only in the coagulation unit). Then, this strategy will be implemented on-line.

The water treatment units contain complex processes. Few researches so far concerned their control or diagnosis. However, front of the more and more necessity of producing a water of constant quality, the producers of potable water become made sensitive to any technique allowing to answer quickly this requirement. For that purpose, the future works aim at establishing a methodology of diagnosis based on the use of wireless sensors and the interpretation of the information obtained on all of the water treatment plant by including the value of coagulant dosage calculated by the neural software sensor.

References

1. SMAPA (2007). Sistema Municipal de Agua Potable y Alcantarillado de Tuxtla, Manual de Procedimientos. Tuxtla Gtz., Chiapas, México.
2. Lamrini, B., Benhammou, A., Le Lann, M-V.: Construction d'un capteur logiciel pour la prédiction de la dose du coagulant: application à une station de traitement d'eau potable. Laboratoire d'automatique, UCA, Marrakech-LAAS/CNRS, Toulouse, France (2004).
3. Valentin, N.: Construction d'un capteur logiciel pour le contrôle automatique du procédé de coagulation traitement d'eau potable. Thèse de doctorat, UTC/L.desEaux/CNRS (2000).
4. Lind, C.: Coagulation Control and Optimization: Part One, Public Works, Oct (1994) 56-57
5. Lind, C.: Coagulation Control and Optimization: Part Two. Public Works, Nov(1994) 32-33
6. K. Baba, I. Enbutu, M. Yoda. Explicit representation of knowledge acquired from plant historical data using neural network. Int. Joint Conf on Neural Networks. Wash, D.C., 1990.
7. A. Mirsepassi, B. Cathers, H.B. Dharmappa. Predicted of chemical dosage in water treatment plants using ANN models. IAWQ Asia-Pacific Reg.Conf. Korea, 1997, 16-561.

8. I. Enbutsu, K. Baba, N. Hara. K. Waseda, A. Nogita. Integration of multi AI paradigms for intelligent operation – fuzzy rule extraction from a neural network. Wat. Res.1998,28,11-12.
9. I. Fletcher, A. Adgar, C.S. Cox, T.J. Boheme. Neural Network applications in the water industry. The Institute of Electrical Engineers IEE pp 16/1-16/6, London, UK, 2001.
10. C.W. Baxter, S.J. Stanley, Q. Zhang, D.W. Smith. Developing artificial neural network process models of water treatment process. Eng. Sci./Rev.gen.sci.env.1(3):pp201-211, 2002.
11. W. Peijing, C. Cox. Study on the application of auto-associative neural network. IEEE ICSP'04 Proceedings, 0-7803-8406-7/04, pp 1570-1573, 2004.
12. H. Hernández De León: Supervision et diagnostic des procédés de production d'éau potable. Thèse de doctorat, INSA Toulouse. France/CNRS (2006)
13. Piera N., Desroches P., Aguilar J.: LAMDA: An incremental Conceptual Clustering System, Report 89420 LAAS-CNRS (1989)
14. Waissmann J.: Construction d'un modèle comportemental pour la supervision de procédés : Application a une station de traitement des eaux, PhD Thesis, LAAS/CNRS, Institut National Polytechnique de Toulouse, France (2000)
15. Aguado J.C.: A Mixed Qualitative-Quantitative Self-Learning Classification Tech applied to Situation Assessment in Industrial Process Control. PhD Thesis. UP Catalunya (1998)
16. Aguilar-Martin, J., López R.: The process of classification and learning the meaning of linguistic descriptors of concepts. Approx Reasoning in Decision Analysis.N.Holland (1982)
17. Waissman J., Aguilar-Martin, J., Dahhou B., Roux G.: Généralisation du degree d'adéq. marginale (DAM) de la classification LAMDA. Soc.Francophone de Classification (1998).
18. Kempowsky, T., Aguilar-Martin, J., Le Lann, M-V., Subias, A.: Learning Methodology of a supervision System using LAMDA Classification Method. LAAS/CNRS.Iberamia (2002)
19. W.S. McCulloch, W. Pitts. A logical calculs of the ideas immanent in nervous activity. Bulletin of Math. 1943, Biophysics, 5, 115-133.
20. D.E. Rumelhart, J.L. McClelland. Parallel distribution processing: exploration in the microstructure of cognition. Cambridge, MA, 1986, MIT Press, 1.
21. W. Duch. Survey of neural transfer functions. Neural Computing Surveys. 1999, 2.
22. Matlab, Neural Network Toolbox, User' Guide, Inc., (2007)
23. Kempowsky, T.: SALSA(Situation Assessment using LAMDA Classification Algorithm). User's Manual. Rapport LAAS/CNRS No. 04160 (2004).