Pergamon

Wat. Sci. Tech. Vol. 36, No. 6-7, pp. 209–217, 1997.

0273–1223/97 \$17·00 + 0·00

PII: S0273-1223(97)00525-8

HYBRID FUZZY NEURAL NETWORK FOR DIAGNOSIS – APPLICATION TO THE ANAEROBIC TREATMENT OF WINE DISTILLERY WASTEWATER IN A FLUIDIZED BED REACTOR

Jean-Philippe Steyer, Damien Rolland,
Jean-Claude Bouvier and René Moletta

Laboratoire de Biotechnologie de l'Environnement, INRA, Avenue des Etangs, 11100 Narbonne, France

ABSTRACT

In this paper, we present a hybrid approach that uses both fuzzy logic and artificial neural networks for on-line detection and analysis of problems occurring in a 120 liter anaerobic digestion fluidized bed reactor for the treatment of wine distillery wastewater. The raw data available on the process (i.e., pH, temperature, recirculation flow rate, input flow rate and gas flow rate) are preprocessed using fuzzy logic to build a vector of features (i.e., a pattern vector). This feature vector is classified into a prespecified category (i.e., a class) which is a state of the system, according to discrimination fuzzy rules. An artificial neural network is then used to classify the process states and to identify the faulty or dangerous ones. This approach was developed to handle in real time problems such as, for example, foam forming, sudden changes in the effluent to be treated (due to a change in concentration), pipe clogging (due to struvite formation) or bad temperature regulation (due to improper setting of the control parameters). © 1997 IAWQ. Published by Elsevier Science Ltd

KEYWORDS

Anaerobic digestion; artificial neural network; diagnosis; fluidized bed; fuzzy logic.

INTRODUCTION

For many years now, research has been carried out on the microbial ecosystem of the anaerobic digestion process, its kinetics, mathematical modeling and on the development of appropriate control strategies. This has allowed one to determine the different parameters involved in the process and the values necessary to perform good operating conditions. In particular, dynamic models have been shown to be very powerful tools to improve monitoring and control of wastewater treatment plants: they can first be used to analyze and to predict the performance of a plant under different operating conditions (Andrews, 1978; Mosey, 1983). They can also help to understand the process from a global point-of-view and to train the process operators (Mc Carthy *et al.*, 1991). But they can also be used directly in the control loop thus improving the running of a process. However, more or less serious breakdowns in industrial applications have been reported owing mainly to the organic overload of various origins. They created some kind of suspicion towards this process and delayed its development. Thus, the importance of implementing efficient diagnosis systems for anaerobic digestion processes is of no doubt. Anaerobic digestion is indeed intrinsically a very unstable process: for example, variations of the input variables (hydraulic flow rate, influent organic load) may easily lead the process to a wash-out of the tank.

A clear need for advanced monitoring and control systems is thus expressed. However, the classical monitoring and control methods are proven to be not very efficient to tide over the internal working and dynamics of the wastewater treatment processes. Indeed, most often there exists only local control actions based on pH and/or temperature regulation for instance which are inefficient to control the key biochemical process variables and to optimize the process operation. An important missing link is clearly the availability of decision support systems capable of improving the relevance of the on-line measurements and of handling the ill-defined available knowledge while controlling the wastewater treatment process.

As a consequence, and because of increasing requirements on reliability and safety of technical processes, many advanced methods have been developed for fault detection (FD) and diagnosis in recent years. However, most of these approaches are based on mathematical models of the process even though global and accurate models are often unavailable in the case of complex systems. This is particularly true for biological wastewater treatment processes that, in addition, are submitted to a wide range of perturbations and thus, to different sources of dysfunctionning. Complementary to advanced control schemes, monitoring systems including diagnosis capabilities can then be of a great help to improve waste removal.

This last decade, fuzzy logic already demonstrated its interest for modeling, control and/or diagnosis of ill-defined complex systems. In particular, fuzzy sets theory allows us to deal with the uncertainty in the recognition process and to elaborate complex decision rules (Frank, 1994). In addition to this approach, the emergence of the neural network paradigm as a powerful tool for learning complex mappings from a set of examples has triggered interest in using neural network models for fault diagnosis (Köppen-Seliger *et al.*, 1995). We thus decided to combine both fuzzy logic and artificial neural networks in a single frame in order to be able to detect any discrepancy occurring on the process.

MATERIALS AND METHODS

The wastewater to be treated is an industrial wine distillery effluent whose COD content is approximately 30 g/l. The process is a fluidized bed reactor of 120 liters working volume. Its architecture is depicted in Figure 1. The reactor is a 0.25 x 3.65 m inox column. Temperature is maintained at 35±2°C by a variable power heat exchanger. A settling device is provided in order to separate the off bioparticles from the liquid phase and to recycle it periodically into the reactor by means of a vortex pump. The support particles are a fine granular material with a specific gravity of 2. The expansion is provided by the recycle flow.

All main parameters (i.e., pH, temperature, liquid and gas flow rates) are measured and stored on a personal computer using advanced monitoring software named Control-BUFFER developed in our laboratory. This software includes several control strategies together with the diagnosis procedure presented here.

A first problem we want to avoid is the inappropriate tuning of control algorithms. We had an example where, due to the bad control that lasted for 15 hours, we were not able to operate the process for the next 75 hours (that is more than 5 times the problem duration). It is obvious that this situation should be avoided at any price. But other problems – much less predictable – may appear even though the control action is appropriate. Among the problems we faced, we can notice:

- pipe clogging (cf. Figure 2): if everything was normal, the valve opening should be constant to keep the flow rate constant. However, because of struvite formation, pipe clogging occurred and valve opening had to increase.
- foam forming (cf. Figure 3): the gas flow rate was constant until foam forming occurred and disturbed the measurement from time T = 10 h to T = 32 h (when antifoam was added).

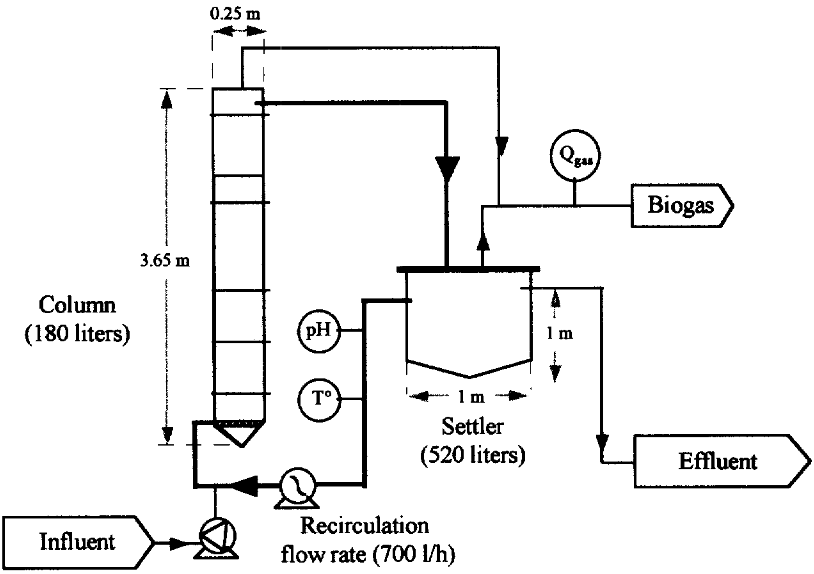

Figure 1. Experimental setting.

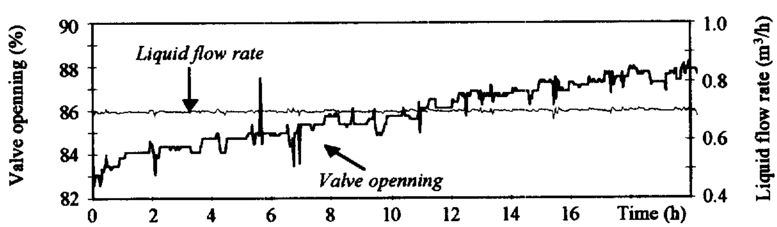

Figure 2. Influence of pipe clogging on the valve opening.

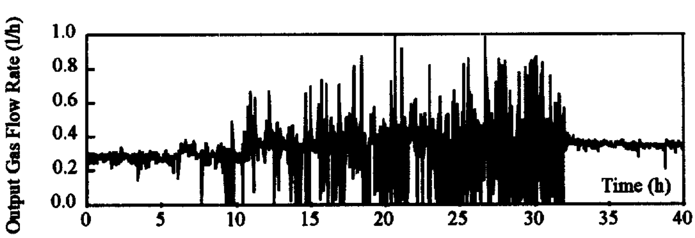

Figure 3. Influence of foam forming on the output gas flow rate (normalized values).

RESULTS AND DISCUSSION

The diagnosis architecture we developed is depicted in Figure 4. It involves several steps. First, every measurement – together with any signal processing and/or calculation related to it (e.g., trend quantification, noise affecting the signal and results from numerical models) – is qualified into a fuzzy value: low, normal or high by comparison to a moving average (cf. Figure 5). This fuzzy qualification is then combined with additional fuzzy rules to determine whether or not the variables are in a faulty situation. Additional information is provided by the membership function of the faulty set since this can be linked to the percentage of fault. Figures 6.a and 6.b illustrate this step. They present the fuzzy diagnosis of two eventual problems related to the analysis of the output gas flow rate shown in Figure 3: *"Diag_Sin"* indicates if the fault detected is a change of the influent substrate concentration and/or feed rate (cf. Figure 6.a) and *"Diag_Foam"* gives information about foam forming. When these variables are equals to 0, then the process is considered in normal operating conditions whereas a value of 1 shows a totally faulty situation (i.e., 100 % of error). Between these values, they indicate the percentage of suspicion that a problem is occurring.

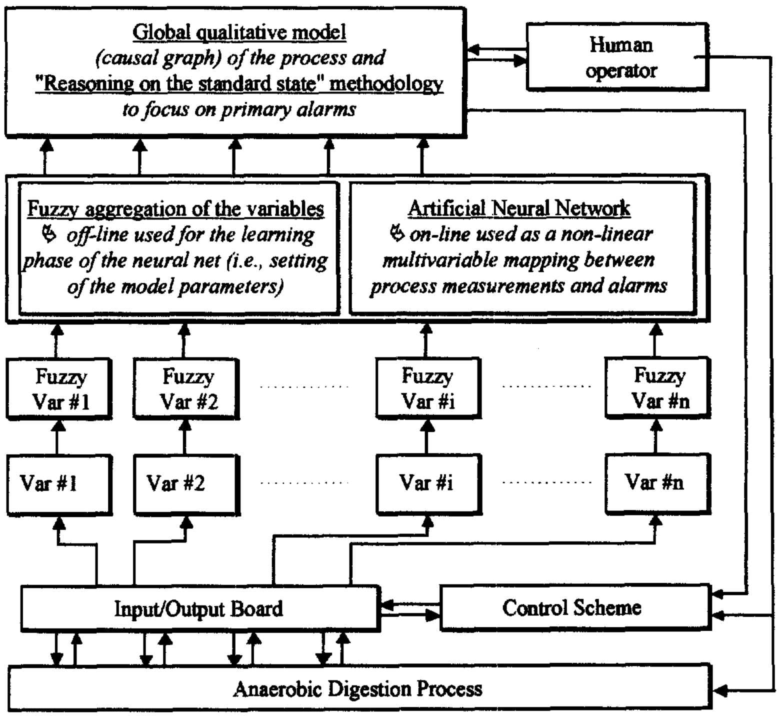

Figure 4. General architecture of our hybrid fuzzy neural network diagnosis scheme.

Then, the variables interactions are accounted for within a fuzzy model of the system. This step provides a vector of feature (i.e., a pattern vector) that is classified into a prespecified category (i.e., a class) according to discrimination fuzzy rules. Alarm filtering (AF) by focusing on the original causes is here the main goal. As a very simple example, this step can be again illustrated using Figures 6.a and 6.b. By classifying *"Diag_Foam"* by a higher order than *"Diag_Sin"*, the diagnosis procedure only focuses on the problem related to foam forming. Another example is presented in Figure 7. In this case, we disturbed on purpose the influent liquid flow rate (Figure 7.a) but we did not indicate it to the FD procedure. Effects of these changes can be seen on the output gas flow rate (Figure 7.b). By just analyzing this evolution, the FD and AF procedures provided the results in Figure 7.c and 7.d. We can see that the diagnosis was well focused on the change of the influent and that foam forming was not suspected.

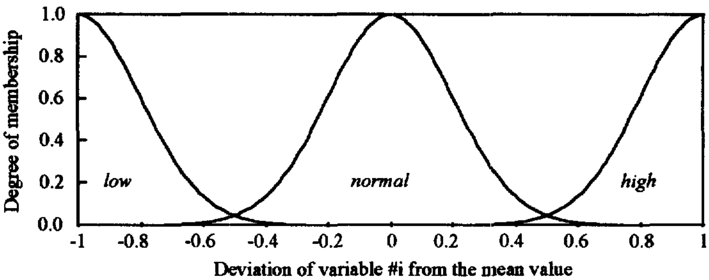

Figure 5. An example of fuzzy qualification of a property of the variable #i.

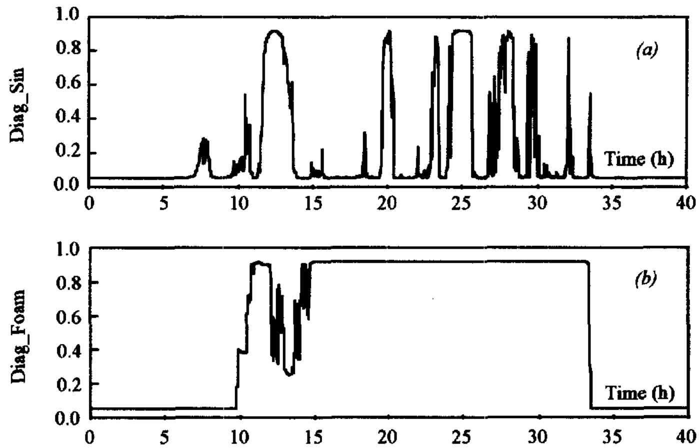

Figure 6. Diagnosis (in percentage of error) on the change of the influent (a) and on the foam forming (b) when analyzing the output gas flow rate presented in Figure 3.

The procedure could have been stopped at the previous step if the process was always in the same operating conditions. Indeed, in such a case, we would not need to adapt the FD thresholds from time to time. However, anaerobic digestion systems – and any wastewater treatment plant in general – is anything but in stable operating conditions. It is obvious when we deal with the start-up of a process: the more the process goes on, the more the biomass is present and active. This has a large influence on the dynamics of the process and thus, thresholds have to be adapted. But thresholds have also to be adapted even though a process is running in the same operating conditions for a long time. Indeed, even though it seems to be in steady state, disturbances like changes in the influent concentration and quantity, together with changes in the pH, the temperature or any variable influence the dynamics of the process. Thus, a particular evolution could lead, one day, to an alarm and, another day, could be judged as a normal situation. When we tackled this problem, we found artificial neural networks (ANN) very efficient to map the measurement changes to the evolving faulty situations. Indeed, they are very efficient tools to handle multivariable non-linear relationships like those we have to deal with. Each variable to be diagnosed is thus related to a specific neural net (i.e., one input and one output). We determined the following ANN architecture to be the most appropriate for our purpose: 2 hidden layers with 4 neurons in the first layer and 2 in the second one, feedforward architecture and training performed using the Levenberg-Marquardt methods. The learning phase of each ANN was done using the fuzzy aggregation of the variables into the fuzzy model. In each case, different faults were used to provide a good mapping capability. Figure 8 shows a comparison of the FD results using fuzzy qualification and ANN when the organic loading rate is changed on purpose (but without indicating it to the FD procedure). We can see that the neural net is more capable of detecting the fault than the fuzzy qualification. It is to be noticed that in this example, the gas flow rate evolution was not part of the training set of the neural network.

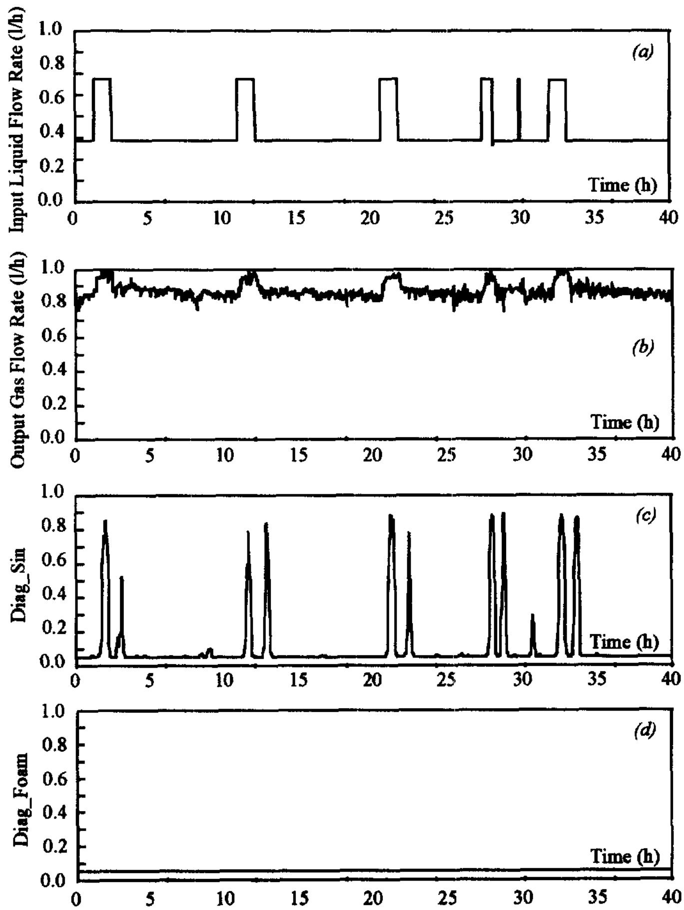

Figure 7. An example of fault detection and alarm filtering when the organic loading rate is changed.

Finally, the last step provides a deeper analysis of the problems. It is different from the second stage that just handled simple relationships between variables and could be viewed as a one-step ahead management of the inter-relationships. Since anaerobic digestion processes are much more complex than straightforward links between variables, we indeed decided to build a general model based on a causal graph representation support. This provides a clear view of causality links and thereby a good support for explanation. In addition, the "biological dimension" inherent in these processes could be expressed within an appropriate formalism, deep enough to get close to the micro-organisms activity but simple enough to be used on-line (Steyer 1991).

We applied this formalism, together with a specific qualitative reasoning technique – the " reasoning on the standard state" (RSS) (Steyer *et al.*, 1996) – and we included statistical tests to automatically manage the cross-relationship between variables. Compared with a classical expert system with IF-THEN rules where every situation has to be accounted for from the beginning (see for example (Barnett *et al.*, 1992)), this approach allows us to build much simpler knowledge bases (KB). An example is presented in Figure 9: the first line is used for the statistical tests and the second line is the KB of the TRC300 variable. Each variable has thus its own KB which is no more than a single line such as, for the variable TRC300: "m *trf* m *qrc* m *cap* m *cal* m".

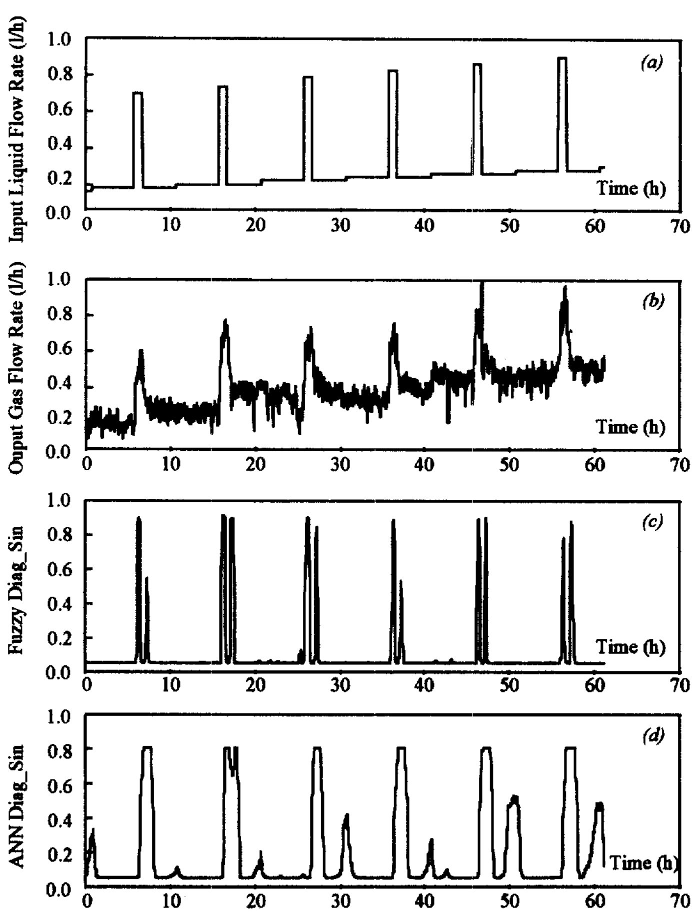

Figure 8. Comparison of fault detection using fuzzy qualification and artificial neural network when the organic loading rate is changed.

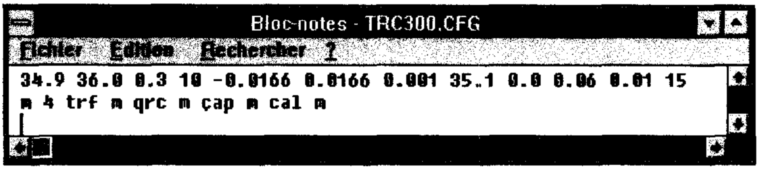

Figure 9. Knowledge base associated to a process variable.

This is to be read as following: the "standard state" (i.e., the optimal state) of TRC300 is "medium" and when its measured state is "medium", *trc, cap* and *cal* should also be "medium". If this is not the case, the RSS methodology automatically generates the list of possible causes of this problem using logical relationships. Then, statistical tests using for example the Table of Student (Schwartz, 1980) are activated to determine the degree of relationship between measurement evolutions in order to focus on the most informative variables. Another important advantage of this approach is the possibility to on-line change – using any text editor – the KB according to the experience gained on the process. In addition, when applied to failure diagnosis and process monitoring, classical expert systems are often based on sets of symptoms of the different faults and/or influences that can affect the system. If, on one hand, the diagnosis problem is simplified, on the other hand, the detection decision may fail if the failures set is not complete. On the opposite, by combining quantitative and qualitative information in a model-based approach, the RSS methodology was demonstrated to be more reliable and more robust.

CONCLUSION

This paper presented an approach combining fuzzy logic, artificial neural networks and a qualitative knowledge-based model for the diagnosis of an anaerobic digestion process. The obtained results demonstrated its ability to handle a large number of problems in a simple frame.

Finally, this approach is to be seen as a complement to control systems. Indeed, control systems are usually dedicated to specific tasks and their choice results from compromise between different objectives. For example, one might want to optimize the starting operation of a plant, to obtain good performances of the process and/or to regulate a substrate concentration at a desired value despite disturbances on the reactor. The control theory can handle separately these objectives but an important missing link is clearly the availability of decision support systems capable of:

- deciding when to switch from one control scheme to another one,
- setting more appropriate values of the tuning parameters of a chosen control scheme,
- improving the relevance of the on-line measurements and handling the process despite problems with a sensor and/or an actuator.

REFERENCES

Andrews, J.F. (1978). The development of a dynamic model and control strategies for the anaerobic digestion process, *Math. Mod. in Wat. Poll. Cont.*, J. Wiley, Chichester, 281.

Barnett, M.W. and Andrews, J.F., (1992). Expert system for anaerobic digestion process operation, *Journal of Environmental Engineering*, **118**(6), 949-963.

Frank, P.M. (1994). Application of fuzzy logic to process supervision and fault diagnosis, *Proc. IFAC Symp. On Fault Detection, Supervision and Safety for Technical Processes*, Espoo, Finland, vol. 2, 531-538.

Köppen-Seliger, B., Frank, P.M. (1995). Fault Detection and isolation in technical processes with neural networks, *Proc. Int. Conf. On Decision and Control*, New Orleans, USA, 2414-2419.

McCarthy, P.L. and Mosey, F.E. (1991). Modelling of anaerobic processes (a discussion of concepts), *Wat. Sci. Tech.*, **24**, 17-33.

Mosey, F.E. (1983). Mathematical modelling of the anaerobic digestion process: Regulatory mechanisms for the formation of short-chain volatile acids from glucose, *Wat. Sci. Tech.*, **15**, 209-232.

Schwartz, D. (1980). *Méthodes statistiques à l'usage des médecins et biologistes*, Collection Statitique en Biologie et Médecine, Flammarion Médecine Science Eds, 311 pages.

Steyer, J-Ph. (1991). *Sur une approche qualitative des systèmes physiques: Aide en temps réel à la conduite de procédés fermentaires*, PhD. Thesis, Université Paul Sabatier, Toulouse, France.

Steyer, J-Ph., Queinnec, I., Capit, F. and Pourciel, J-B. (1996). Qualitative rules as a way to handle the biological state of a fermentation process: An industrial application, *Journal Européen des Systèmes Automatisés JESA-RAIRO-APII*, **30**(2/3), 381-398.