

# Fault Diagnosis in a Yeast Fermentation Bioreactor by Genetic Fuzzy System

Tayyebi, Shokoufe; Shahrokhi, Mohammad\*<sup>+</sup>; Bozorgmehry Boozarjomehry, Ramin

Faculty of Chemical and Petroleum Engineering, Sharif University of Technology,

P.O. Box 11365-9465 Tehran, I.R. IRAN

**ABSTRACT:** In this paper, the fuzzy system has been used for fault detection and diagnosis of a yeast fermentation bioreactor based on measurements corrupted by noise. In one case, parameters of membership functions are selected in a conventional manner. In another case, using certainty factors between normal and faulty conditions the optimal values of these parameters have been obtained through the genetic algorithm. These two cases are compared based on their performances in fault diagnosis of a yeast fermentation bioreactor for three different conditions. The simulation results indicate that the fuzzy-genetic system is superior in multiple fault detection for the conditions where the minimum and maximum deviations from normal conditions occur in the process variables.

**KEY WORDS:** Fault detection and diagnosis, Fermentation bioreactor, Fuzzy logic, Genetic algorithm.

## INTRODUCION

A failure can be described as diversity of malfunction in the real plant which can be caused due to instruments failures, disturbances and plant parameters uncertainties. The abnormal conditions in a plant can result in financial losses. Therefore, in the chemical processes, fault diagnosis has been the focal point of many researches and various fault detection and diagnosis strategies have been presented in the literature. The fault diagnostic systems should possess desirable characteristics such as quick detection, isolability, robustness and multiple fault identifiability [1].

The fault detection strategies can be classified in two groups based on the type of knowledge and information

being used. In the first group, the mathematical model of the system is used for fault detection like Signed Direct Graph (SDG), Fault tree, Extended Kalman Filter (EKF). For example, Cui *et al.* [2] presented a model based online fault detection and diagnosis strategy for centrifugal chiller systems. Sotomayor *et al.* [3] utilized unknown input observer methods for designing fault diagnosis systems for a styrene polymerization reactor and the FCC reactor-regenerator system. Nelly Olivier-Maget *et al.* [4] presented a method for the fault detection based on the comparison between the reference model evolution and the real system generated by the extended Kalman filter for the industrial plant.

---

\* To whom correspondence should be addressed.

+ E-mail: shahrokhi@sharif.edu

1021-9986/10/3/63

12/\$/3.20

In most cases, finding rigorous mathematical model of the process is very difficult or even impossible. Therefore, in order to overcome this problem, model independent strategies have been used [5]. Among these methods statistical based techniques and intelligent strategies have been used frequently. *Genovesi et al.* [6] presented a fuzzy logic based diagnosis system and implemented it, for the fault diagnosis of a wastewater treatment plant. *Eslamloueyan et al.* [7] designed an efficient probabilistic neural network for fault diagnosis of nonlinear processes operating at multiple operating regions. *He et al.* [8] proposed a new fault diagnosis approach with variable weighted kernel Fisher discriminant analysis for the Tennessee Eastman process. *Detroja et al.* [9] presented an approach based on the Correspondence Analysis (CA) for fault diagnosis systems. *Cen Nan et al.* [10] proposed a knowledge-based fault diagnosis method using the valuable knowledge from the experts and operators for a micro steam power.

In this work, an intelligent strategy (fuzzy logic) is proposed for fault detection and diagnosis of a yeast fermentation bioreactor. First a conventional fuzzy system has been designed for fault diagnosis. Second, genetic algorithm is used to select the appropriate parameters of membership functions for optimizing the fault diagnosis designs. The performances of the proposed strategies are evaluated for two groups of faults. The first group has the minimum deviation from the normal condition while the second group has the maximum deviation.

## FAULT DIAGNOSIS BASED ON FUZZY LOGIC

The importance of Fuzzy Logic (FL) derives from the fact that most modes of human reasoning and especially common sense reasoning are approximative in nature [11]. Fuzzy logic has many applications in medicine [12], control [13], modeling [14], fault detection, diagnosis [15] and so on. Fuzzy systems can handle problems with imprecise and incomplete data, and it can also approximate various types of nonlinear functions. Furthermore, the implementation of FL is flexible and easy. The fuzzy fault diagnosis consists of the following three steps:

1. Fuzzification: The fuzzifier converts the crisp input parameters into linguistic variables. The membership function transforms the degree of fuzziness into the normalized interval [0,1]. Usually, each linguistic

Table 1: The MIMO form of linguistic rules.

If	Premise	Then	Consequent
	$u_1$ is $\tilde{A}_1^{s_1}$ and $u_2$ is $\tilde{A}_2^{s_2}$ and ... and $u_m$ is $\tilde{A}_m^{s_m}$		$y_1$ is $\tilde{B}_1^r$
	$u_1$ is $\tilde{A}_1^{s_1}$ and $u_2$ is $\tilde{A}_2^{s_2}$ and ... and $u_m$ is $\tilde{A}_m^{s_m}$		$y_2$ is $\tilde{B}_2^r$
	.		
	.		
	$u_1$ is $\tilde{A}_1^{s_1}$ and $u_2$ is $\tilde{A}_2^{s_2}$ and ... and $u_m$ is $\tilde{A}_m^{s_m}$		$y_n$ is $\tilde{B}_n^r$

variable has several fuzzy sets. A fuzzy set is represented by a set of pairs  $M_i/C_i$ , where  $M_i$  is the membership function for the linguistic value  $C_i$ .

2. Fuzzy rule base: The rule base consists of a set of linguistic IF-THEN rules. Fuzzy rules link the input variables of fuzzy system to its output variables. In fault diagnosis based on fuzzy logic, the fuzzy relations between faults and symptoms are assumed to be one to many (i.e. one fault may cause several symptoms). The number and complexity of rules depend on the number of input parameters that are to be processed and the number of fuzzy sets associated with each parameter. The MIMO rule with  $m$  input and  $n$  output is shown in Table 1. In this table  $u$  and  $y$  are input and output of linguistic variables, respectively.  $\tilde{A}_i^j$  denote the  $j^{\text{th}}$  linguistic value of the linguistic variable  $u_i$  and  $\tilde{B}_i^j$  indicate the  $j^{\text{th}}$  linguistic value of the linguistic variable  $y_i$ .

3. Inference engine: To define the relationship between fuzzy sets of inputs and outputs, fuzzy inference should be established. For example:

If  $\text{Sym}_1$  is  $S_{1,n}$  AND  $\text{Sym}_2$  is  $S_{2,n}$  AND ... AND  $\text{Sym}_m$  is  $S_{m,n}$  Then  $F_1$  is  $H_1$  where  $\text{Sym}_i$  ( $i = 1 \dots m$ ) is the vector of fuzzy input variables (symptoms) and  $F_j$  ( $j = 1 \dots n$ ) is the fuzzy output variables (Faults).  $S_{i,j}$  is the input linguistic value relevant to  $j^{\text{th}}$  output and  $H_j$  is the output linguistic value. Then, the degree of membership of each linguistic value of the output will be expressed as follows:

$$\mu_{H_h} = \prod \mu_{S_{i,j}} \quad (1)$$

where  $\mu$  is the final membership value resulted by fuzzy logic. The schematic of a fuzzy fault diagnosis system is shown in Fig. 1.

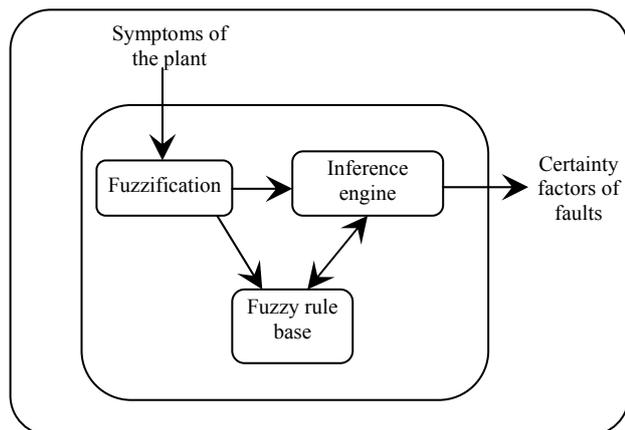


Fig. 1: Schematic diagram of a fuzzy fault detection system.

The input and output membership functions of a fuzzy fault diagnosing system can be triangular functions, trapezoidal functions or Gaussian functions. The most flexible and accurate type of membership function is the Gaussian function which is obtained at the expense of more computational load. They are not usually used for systems with lots of inputs and outputs. However, since the number of inputs and outputs for the system under consideration used is not large, in order to have an accurate and yet flexible system the Gaussian membership functions are used. Five applied linguistic fuzzy sets for the input and output variables are NL (Negative Large), NS (Negative Small), Z (Zero), PS (Positive Small) and PL (Positive Large). The functions of fuzzy sets can be determined according to the following equations:

For fuzzy set NL:

$$\mu_N = \begin{cases} = 1 & \text{if } \bar{x} \leq C_N \\ = \exp(-.5((\bar{x} - C_N)/\sigma_N)^2) & \text{else} \end{cases} \quad (2)$$

For fuzzy set PL:

$$\mu_P = \begin{cases} = 1 & \text{if } \bar{x} \geq C_P \\ = \exp(-.5((\bar{x} - C_P)/\sigma_P)^2) & \text{else} \end{cases} \quad (3)$$

For fuzzy set Z, NS and PS:

$$\mu_Z = \exp(-.5((\bar{x} - C_Z)/\sigma_Z)^2) \quad (4)$$

Where  $\bar{x}$  is the deviation of the normal variable.  $\mu$  is the degree of the membership function and  $c, \sigma$  are the parameters of the membership function.

The corresponding membership function can be constructed according to Fig. 2.

It should be mentioned that, using an algorithmic

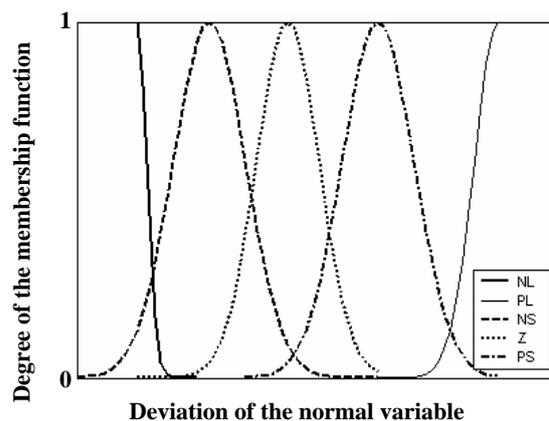


Fig. 2: Schematic of the applied fuzzy sets.

solution may improve the results of the fuzzy diagnosis system. As too many parameters of fuzzy sets exist and their extreme values are wide-ranged, an optimization method should be used in order to obtain the appropriate values for the parameters of each fuzzy set. In the present work, a genetic algorithm has been used for the design of the parameters of fault diagnosis system which will be described in the following section.

## THE GENETIC FUZZY DIAGNOSIS SYSTEM

Genetic algorithm (GA) can be considered as one of the available approaches for global optimization which is inspired from the evolutionary mechanisms that exist in the Mother Nature. In the binary GA the decision variables whose optimum values are being sought are encoded in a binary string called chromosome. Each part of this binary string (called genotype) corresponds to a specific decision variable (called phenotype). The algorithm uses a set of chromosomes (called population) rather than a single one which increases the probability of reaching the global optimum. Fitness function which is somehow related to the objective function represents the suitability of each chromosome. The cycle of the 'fitness' evaluation of all chromosomes and obtaining the new population is called a generation. The chromosomes of the new generation are obtained based on genetic operators. The selection operator chooses two chromosomes from a set of chromosomes (called mating pool) in which the chromosomes with higher fitness values have more copies than those with poor fitness. This approach is used to make sure that chromosomes with better fitness have priority to participate in the production of new

chromosomes. The selected chromosomes produce two new chromosomes called offspring through the following procedures:

(a) Mating the selected two chromosomes from mating pool using the crossover operator.

(b) Random change of the newly produced chromosome using the mutation operator.

The crossover operator consists of the information exchange between parent chromosomes in order to generate two offspring. The mutation operator makes random changes on the selected genes of each offspring. The algorithm goes through numerous generations in which the chromosomes are evolved to increase their fitness until the termination criterion is met [16]. There are various termination criteria in the literature, the most frequently used approaches are as follows:

- 1) Reaching the maximum number of generations
- 2) Getting a small difference between the average and maximum fitness of the population
- 3) Getting a population in which the percent of similarity between the chromosomes is beyond a threshold

As can be seen in equations 2, 3 and 4, the number of the Gaussian function parameters in each fuzzy set is two. Therefore, the number of adjustable variables is calculated by the following equation:

$$n_{AG} = 2 * n_{FS} * n_{MV} \quad (5)$$

where  $n_{AG}$  is the number of adjustable variables,  $n_{FS}$  is the number of fuzzy sets for each variable and  $n_{MV}$  is the number of measured variables. Consequently, the  $n_{AG}$  parameters should be determined by genetic algorithm.

In this paper, the genetic algorithm has been used to optimize membership function parameters of fuzzy diagnosis system. The structure of the chromosome used in the Genetic Algorithm has been designed such that it contains all the parameters of the fuzzy diagnosing System. This structure has been shown in Fig. 3. The fitness function of the genetic fuzzy diagnosis system is the difference of certainty factors between normal and faulty conditions. If the parameters of fuzzy sets which are decrypted from the chromosome whose fitness is evaluated, are appropriate and lead to an efficient and accurate fuzzy diagnosis system, the fitness value of the chromosome will be almost one, otherwise its fitness would be a smaller number. Hence, the best chromosome

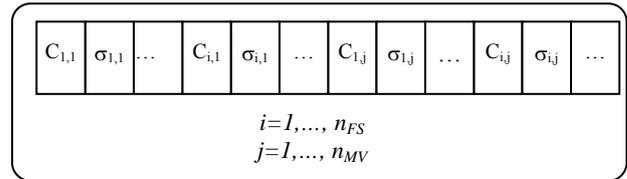


Fig. 3: The structure of the chromosome used in the GA.

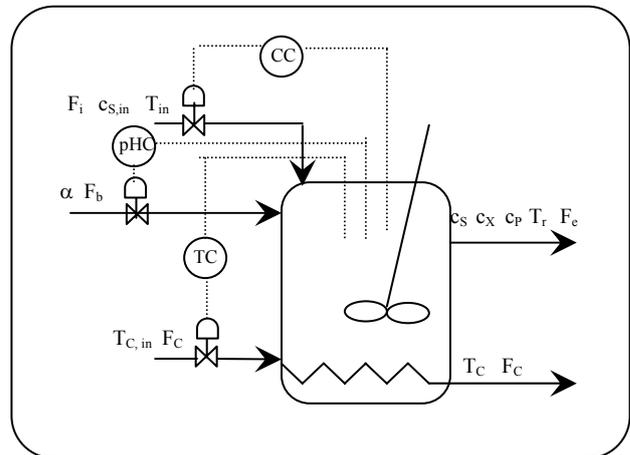


Fig. 4: Schematic diagram of the fermentation reactor.

of each generation represents the parameters of the fuzzy diagnosing system which gets more accurate and efficient. At the time when the similarity ratio of the chromosomes in each generation is beyond a certain threshold (about 96%), the GA has converged to the optimum set of parameters representing the design of appropriate fuzzy diagnosing system.

An initial population of Chromosomes is randomly generated. The population in each generation is taken as 100 and crossover probability and mutation probabilities are chosen to be 0.7 and 0.1 respectively. The number of bits used to span the interval of each parameter is assumed to be 10.

## CASE STUDY

The considered case study is a yeast fermentation bioreactor that is shown schematically in Fig. 4. This process is one of the most frequently used biochemical processes. The process model can be obtained using material and energy balances as presented by Nagy [17]:

Material balance:

$$\frac{dc_X}{dt} = \mu_X c_X \frac{c_S}{K_S + c_S} e^{-K_P c_P} - \frac{F_e}{V} c_X \quad (6)$$

$$\frac{dc_S}{dt} = -\frac{1}{R_{SX}} \mu_X c_X \frac{c_S}{K_S + c_S} e^{-K_P c_P} - \quad (7)$$

$$\frac{1}{R_{SP}} \mu_P c_X \frac{c_S}{K_{S1} + c_S} e^{-K_{P1} c_P} + \frac{F_1}{V} c_{S,in} - \frac{F_e}{V} c_P$$

$$\frac{dc_P}{dt} = \mu_P c_X \frac{c_S}{K_{S1} + c_S} e^{-K_{P1} c_P} - \frac{F_e}{V} c_P \quad (8)$$

where  $c_x$ ,  $c_s$  and  $c_p$  are concentrations of biomass, substrate and product respectively.

The concentration of the dissolved oxygen in the reactor is evaluated from the following equation:

$$\frac{dc_{O_2}}{dt} = (k_1 a)(c_{O_2}^* - c_{O_2}) - r_{O_2} \quad (9)$$

where rate of oxygen consumption is given by:

$$r_{O_2} = \mu_{O_2} \frac{1}{Y_{O_2}} c_X \frac{c_{O_2}}{K_{O_2} + c_{O_2}} \quad (10)$$

Energy balance:

$$\frac{dT_r}{dt} = \frac{F_1}{V} (T_{in} + 273) - \frac{F_e}{V} (T_r + 273) + \quad (11)$$

$$\frac{r_{O_2} \Delta H_r}{32 \rho_r C_{heat,r}} - \frac{K_T A_T (T_r - T_{ag})}{V \rho_C C_{heat,C}}$$

$$\frac{dT_C}{dt} = \frac{F_C}{V_j} (T_{C,in} - T_C) + \frac{K_T A_T (T_r - T_C)}{V \rho_C C_{Heat,C}} \quad (12)$$

where  $T_r$  and  $T_c$  are reactor and jacket temperatures.

To model pH variation, it is assumed that the microorganisms produce acid (HA) and a base (BOH) is used for neutralization and pH control.  $x_1$  and  $x_2$  are defined as given below:

$$x_1 = [HA] + [A^-] \quad (13)$$

$$x_2 = [BON] + [B^+] \quad (14)$$

Material balance on A and B yields:

$$\frac{dx_1}{dt} = -\frac{F_e}{V} x_1 + x_{1,gen} \quad (15)$$

$$\frac{dx_2}{dt} = \frac{F_b}{V} \alpha - \frac{(F_b + F_e)}{V} x_2 \quad (16)$$

Where  $F_b$  is the base flow rate and  $x_{1,gen}$  is given by:

$$x_{1,gen} = \frac{H_{gen}^2}{k_a} + H_{gen} \quad (17)$$

$H_{gen}$  is the rate of hydrogen production by microorganism obtained from the following equation:

$$H_{gen} = \frac{\mu_X c_X}{Y_{X/H}} \quad (18)$$

The reactor pH is related to  $x_1$  and  $x_2$  through neutrality condition as given below:

$$-\frac{x_1}{1 + 10^{(PK_a - PH)}} + \frac{10^{PH}}{10^{-PH} + 10^{PK_\beta - PK_w}} x_2 + \quad (19)$$

$$10^{-PH} - k_w 10^{PH} = 0$$

For detailed information the readers are referred to articles in which the details of the modeling approach have been expressed [17, 18].

Model parameters are presented in Table 2. For this system, three PI controllers are used to control temperature, product (ethanol) concentration and pH of the reactor by manipulating coolant flow rate, flow of the substrate entering the reactor and the input flow rate of the base. The measurements singles in the process have been all corrupted by zero mean white noises. Furthermore, it is assumed that the controllers are always working appropriately and no fault occurs in their operation.

Faults are divided in two groups based on the level of deviations from their corresponding normal values.

The 1<sup>st</sup> group:

Minimum deviations from normal conditions:

F<sub>1</sub> 1% decrease in the feed temperature ( $T_{in}$ )

F<sub>2</sub> 1% increase in the glucose concentration in the feed flow ( $C_{s, in}$ )

F<sub>3</sub> 5% increase in the coolant input temperature ( $T_{c, in}$ )

F<sub>4</sub> 0.5% decrease in the heat transfer coefficient ( $K_T$ )

The 2<sup>nd</sup> group:

Maximum deviations from normal conditions:

F<sub>1</sub> 20% decrease in the feed temperature ( $T_{in}$ )

F<sub>2</sub> 20% increase in the glucose concentration in the feed flow ( $C_{s, in}$ )

F<sub>3</sub> 15% increase in the coolant input temperature ( $T_{c, in}$ )

F<sub>4</sub> 20% decrease in the heat transfer coefficient ( $K_T$ )

It is assumed that the following variables are measured for process fault diagnosis:

1. Coolant flow rate ( $F_c$ )

2. Feed flow rate ( $F_i$ )

3. Base flow rate ( $F_b$ )

Table 2: Parameters of the process model.

$A_1 = 9.5 \times 10^8$	$H_{CO_2} = 0.485$	$m_{CaCO_3} = 100 \text{ g}$
$A_2 = 2.55 \times 10^{33}$	$H_{H_2O} = 0.941$	$m_{MgCl_2} = 100 \text{ g}$
$C_{heat,ag} = 4.18 \text{ J.gr}^{-1}.K^{-1}$	$K_{O_2} = 8.86 \text{ mgr/l}$	$R_{SX} = 0.607$
$C_{heat,r} = 4.18 \text{ J.gr}^{-1}.K^{-1}$	$K_{P1} = 0.07 \text{ g/l}$	$T_{in} = 25 \text{ }^\circ\text{C}$
$C_{S,in} = 60 \text{ g/l}$	$K_P = 0.139 \text{ g/l}$	$T_{in,C} = 15 \text{ }^\circ\text{C}$
$E_{a1} = 55000 \text{ j/mol}$	$K_S = 1.030 \text{ g/l}$	$V = 1000$
$E_{a2} = 220000 \text{ j/mol}$	$K_{S1} = 1.030 \text{ g/l}$	$V_j = 50 \text{ l}$
$H_{Na} = -0.550$	$K_T = 3.6 \times 10^5 \text{ J.h}^{-1}.m^{-1}.K^{-1}$	$Y_{O_2} = 0.970 \text{ mg/mg}$
$H_{Ca} = -0.303$	$K_a = 8.32 \times 10^{-4} \text{ mol/l}$	$\Delta H_r = 518 \text{ kJ/mol}$
$H_{Mg} = -0.314$	$K_b = 1.78 \times 10^{-5} \text{ mol/l}$	$\mu_{O_2} = 500 \text{ h}^{-1}$
$H_H = -0.774$	$K_w = 10^{-14} \text{ mol/l}$	$\mu_p = 1.790 \text{ h}^{-1}$
$H_{Cl} = 0.844$	$m_{NaCl} = 500 \text{ g}$	$\rho_r = 1080 \text{ g/l}$
$(k_1a)_0 = 38 \text{ h}^{-1}$	$Y_{x/H} = 9330$	$\rho_c = 1000 \text{ g/l}$

Table 3: Results of 1<sup>st</sup> group fault detection by conventional fuzzy system for single fault using three measured variables.

Single fault	Certainty factors	Final results
Normal	[0.00 0.00 0.00 0.00]	Normal
F <sub>1</sub>	[0.74 0.00 0.67 0.72]	F <sub>1</sub> ,F <sub>3</sub> ,F <sub>4</sub>
F <sub>2</sub>	[0.01 0.80 0.00 0.01]	F <sub>2</sub>
F <sub>3</sub>	[0.65 0.00 0.69 0.67]	F <sub>1</sub> ,F <sub>3</sub> ,F <sub>4</sub>
F <sub>4</sub>	[0.64 0.00 0.66 0.66]	F <sub>1</sub> ,F <sub>3</sub> ,F <sub>4</sub>

### The 1<sup>st</sup> group

The results of conventional diagnosing fuzzy system and fuzzy-GA system for single faults are presented in Tables 3 and 4.

The results indicate that the two proposed strategies can't diagnose the plant faults. The reason can be explained as follows:

Variations of  $T_{in}$ ,  $T_{c,in}$  or  $K_T$  have similar effects on  $F_c$  and therefore their corresponding faults cannot be distinguished and hence cannot be detected. This is mainly due to the fact that the system responds similarly to these three different inputs.

To solve this problem, temperature of outlet cooling water has been added to the measured variables. The rules which describe the faults of the plant can be stated as given in Table 5.

The Single, double, triple, quadruple faults diagnosis results based on conventional fuzzy and fuzzy-GA system

Table 4: Results of 1<sup>st</sup> group fault detection by fuzzy-GA system for single fault using three measured variables.

Single fault	Certainty factors	Final results
Normal	[0.00 0.00 0.00 0.00]	Normal
F <sub>1</sub>	[0.86 0.00 0.68 0.73]	F <sub>1</sub> ,F <sub>3</sub> ,F <sub>4</sub>
F <sub>2</sub>	[0.00 0.86 0.00 0.01]	F <sub>2</sub>
F <sub>3</sub>	[0.43 0.00 0.77 0.64]	F <sub>3</sub> ,F <sub>4</sub>
F <sub>4</sub>	[0.64 0.00 0.72 0.73]	F <sub>1</sub> ,F <sub>3</sub> ,F <sub>4</sub>

are given in Tables 6 and 7. As can be seen all faults have been detected correctly. It is due to the fact that after the outlet temperature of the cooling water is measured, there is a unique set of outputs for each set of inputs. Therefore, the proposed strategies can identify the faulty variables successfully.

Figs. 5 and 6 show the average fitness and similarity ratio in each generation of GA until the algorithm converges to its final results.

The selected parameters of each fuzzy set are listed in Table 8 and also the optimized membership functions of input variables are presented in Table 9.

### The 2<sup>nd</sup> group

The diagnosis results based on conventional fuzzy and fuzzy-GA when  $T_c$  is added to the measurement variables are given in Tables 10 and 11.

**Table 5: Rules that applied for fuzzy diagnosis bioreactor system.**

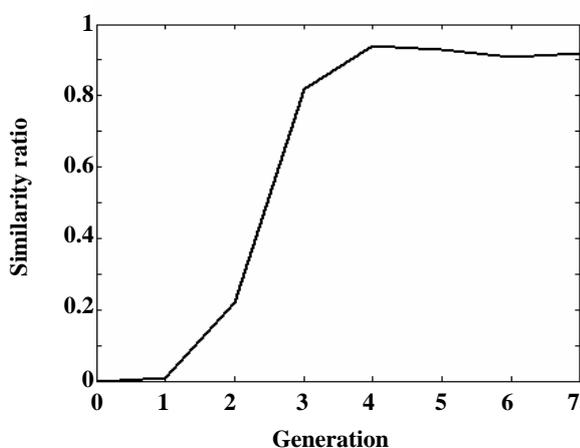
Number of rule	Rule
1	IF $F_c$ is NL And $F_i$ is Z And $F_b$ is Z And $T_c$ is PL Then $F_1$ is NL
2	IF $F_c$ is PL And $F_i$ is Z And $F_b$ is Z And $T_c$ is NL Then $F_1$ is PL
3	IF $F_c$ is PL And $F_i$ is NL And $F_b$ is PL And $T_c$ is NL Then $F_2$ is NL
4	IF $F_c$ is NL And $F_i$ is PL And $F_b$ is NL And $T_c$ is PL Then $F_2$ is PL
5	IF $F_c$ is PL And $F_i$ is Z And $F_b$ is Z And $T_c$ is Z Then $F_3$ is PL
6	IF $F_c$ is NL And $F_i$ is Z And $F_b$ is Z And $T_c$ is Z Then $F_3$ is NL
7	IF $F_c$ is PL And $F_i$ is Z And $F_b$ is Z And $T_c$ is NL Then $F_4$ is NL
8	IF $F_c$ is NL And $F_i$ is Z And $F_b$ is Z And $T_c$ is PL Then $F_4$ is PL
9	IF $F_c$ is NS And $F_i$ is Z And $F_b$ is Z And $T_c$ is PS Then $F_1$ is NS
10	IF $F_c$ is PS And $F_i$ is Z And $F_b$ is Z And $T_c$ is NS Then $F_1$ is PS
11	IF $F_c$ is NS And $F_i$ is NS And $F_b$ is PS And $T_c$ is NS Then $F_2$ is NS
12	IF $F_c$ is PS And $F_i$ is PS And $F_b$ is NS And $T_c$ is PS Then $F_2$ is PS
13	IF $F_c$ is PS And $F_i$ is Z And $F_b$ is Z And $T_c$ is Z Then $F_3$ is PS
14	IF $F_c$ is NS And $F_i$ is Z And $F_b$ is Z And $T_c$ is Z Then $F_3$ is NS
15	IF $F_c$ is PS And $F_i$ is Z And $F_b$ is Z And $T_c$ is NS Then $F_4$ is NS
16	IF $F_c$ is NS And $F_i$ is Z And $F_b$ is Z And $T_c$ is PS Then $F_4$ is PS

**Table 6: Results of 1<sup>st</sup> group fault detection by conventional fuzzy system when  $T_c$  is added to the measured variables.**

Faults	Certainty factors	Final results
Normal	[0.00 0.00 0.00 0.00]	Normal
$F_1$	[0.82 0.02 0.01 0.04]	$F_1$
$F_2$	[0.00 0.76 0.00 0.00]	$F_2$
$F_3$	[0.00 0.03 0.85 0.01]	$F_3$
$F_4$	[0.00 0.03 0.04 0.83]	$F_4$
$F_1, F_2$	[0.81 0.80 0.01 0.06]	$F_1, F_2$
$F_1, F_3$	[0.87 0.00 0.84 0.03]	$F_1, F_3$
$F_1, F_4$	[0.79 0.00 0.03 0.75]	$F_1, F_4$
$F_2, F_3$	[0.04 0.89 0.83 0.01]	$F_2, F_3$
$F_2, F_4$	[0.03 0.86 0.00 0.89]	$F_2, F_4$
$F_3, F_4$	[0.04 0.00 0.82 0.75]	$F_3, F_4$
$F_1, F_2, F_3$	[0.83 0.81 0.85 0.06]	$F_1, F_2, F_3$
$F_1, F_2, F_4$	[0.85 0.84 0.04 0.79]	$F_1, F_2, F_4$
$F_2, F_3, F_4$	[0.03 0.87 0.86 0.85]	$F_2, F_3, F_4$
$F_1, F_3, F_4$	[0.84 0.00 0.83 0.81]	$F_1, F_3, F_4$
$F_1, F_2, F_3, F_4$	[0.78 0.83 0.86 0.80]	$F_1, F_2, F_3, F_4$

**Table 7: Results of 1<sup>st</sup> group fault detection by fuzzy-GA system when  $T_c$  is added to the measured variables.**

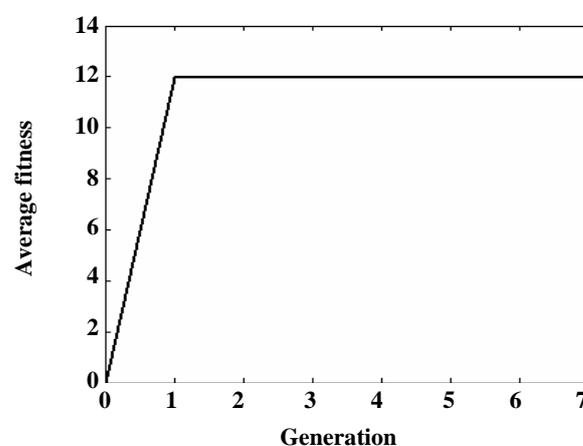
Faults	Certainty factors	Final results
Normal	[0.00 0.00 0.00 0.00]	Normal
F <sub>1</sub>	[0.98 0.00 0.01 0.00]	F <sub>1</sub>
F <sub>2</sub>	[0.00 0.97 0.00 0.00]	F <sub>2</sub>
F <sub>3</sub>	[0.00 0.00 0.99 0.01]	F <sub>3</sub>
F <sub>4</sub>	[0.01 0.00 0.00 0.95]	F <sub>4</sub>
F <sub>1</sub> ,F <sub>2</sub>	[0.95 0.92 0.01 0.00]	F <sub>1</sub> ,F <sub>2</sub>
F <sub>1</sub> ,F <sub>3</sub>	[0.95 0.00 0.93 0.02]	F <sub>1</sub> ,F <sub>3</sub>
F <sub>1</sub> ,F <sub>4</sub>	[0.97 0.00 0.01 0.94]	F <sub>1</sub> ,F <sub>4</sub>
F <sub>2</sub> ,F <sub>3</sub>	[0.00 0.96 0.99 0.02]	F <sub>2</sub> ,F <sub>3</sub>
F <sub>2</sub> ,F <sub>4</sub>	[0.00 0.98 0.00 0.96]	F <sub>2</sub> ,F <sub>4</sub>
F <sub>3</sub> ,F <sub>4</sub>	[0.00 0.00 0.95 0.98]	F <sub>3</sub> ,F <sub>4</sub>
F <sub>1</sub> ,F <sub>2</sub> ,F <sub>3</sub>	[0.97 0.96 0.95 0.01]	F <sub>1</sub> ,F <sub>2</sub> ,F <sub>3</sub>
F <sub>1</sub> ,F <sub>2</sub> ,F <sub>4</sub>	[0.95 0.93 0.00 0.98]	F <sub>1</sub> ,F <sub>2</sub> ,F <sub>4</sub>
F <sub>2</sub> ,F <sub>3</sub> ,F <sub>4</sub>	[0.02 0.99 0.98 0.97]	F <sub>2</sub> ,F <sub>3</sub> ,F <sub>4</sub>
F <sub>1</sub> ,F <sub>3</sub> ,F <sub>4</sub>	[0.97 0.00 0.99 0.98]	F <sub>1</sub> ,F <sub>3</sub> ,F <sub>4</sub>
F <sub>1</sub> ,F <sub>2</sub> ,F <sub>3</sub> ,F <sub>4</sub>	[0.97 0.96 0.95 0.98]	F <sub>1</sub> ,F <sub>2</sub> ,F <sub>3</sub> ,F <sub>4</sub>

**Fig. 5: The similarity ratio versus generation in the 1<sup>st</sup> group.**

The results show that the proposed methods are able to classify four single faults, six double faults, four triple faults and quadruple faults correctly. As can be seen the genetic fuzzy system diagnoses the faults more accurately than the conventional fuzzy system.

## CONCLUSIONS

In this paper, two fuzzy diagnostic systems are

**Fig. 6: The average fitness versus generation in the 1<sup>st</sup> group.**

proposed for a yeast fermentation reactor. The first one is a conventional fuzzy system, while the second one is a fuzzy system whose membership function parameters are optimized by Genetic Algorithm in order to improve its performance. It is shown that with three measured variables, the system is not detectable; therefore one more measurement has been added to make the system detectable. To test the performance of the proposed

**Table 8: The selected parameters of membership functions for the input variables.**

Fuzzy set	The measured variable ( $F_c$ )		The measured variable ( $F_i$ )		The measured variable ( $F_b$ )		The measured variable ( $T_c$ )	
	C	$\sigma$	C	$\sigma$	C	$\sigma$	C	$\sigma$
NS	-5	50	-0.3	10	-3.5e-4	40	-0.5	10
NL	-11	20	-1	50	-5.7e-4	5	-1.8	20
Z	-1	5	0	8	0	12	0.25	3
PS	3	20	0.5	50	4e-5	5	0.9	30
PL	8	30	1	10	8e-5	10	1.8	15

**Table 9: The optimized parameters of membership functions for the input variables.**

Fuzzy set	The measured variable ( $F_c$ )		The measured variable ( $F_i$ )		The measured variable ( $F_b$ )		The measured variable ( $T_c$ )	
	C	$\sigma$	C	$\sigma$	C	$\sigma$	C	$\sigma$
NS	-0.55	42.66	-0.82	98.29	-5.1e-4	141.41	-0.49	39.31
NL	-7.78	36.31	-0.93	9.30	-5.7e-4	63.48	-1.93	124.74
Z	1.12	33.19	-0.52	29.11	-3.1e-4	8.28	1.02	60.13
PS	1.67	105.87	0.25	129.32	1.20e-5	19.64	1.34	130.48
PL	2.84	87.08	0.87	23.28	3.97e-5	39.45	1.56	124.07

**Table 10: Results of 2<sup>nd</sup> group fault detection by conventional fuzzy system when  $T_c$  is added to the measured variables.**

Faults	Certainty factors	Final results
Normal	[0.00 0.00 0.00 0.00]	Normal
$F_1$	[0.85 0.00 0.04 0.06]	$F_1$
$F_2$	[0.05 0.84 0.07 0.03]	$F_2$
$F_3$	[0.06 0.00 0.80 0.05]	$F_3$
$F_4$	[0.06 0.00 0.07 0.78]	$F_4$
$F_1, F_2$	[0.81 0.79 0.06 0.08]	$F_1, F_2$
$F_1, F_3$	[0.85 0.00 0.81 0.00]	$F_1, F_3$
$F_1, F_4$	[0.82 0.00 0.05 0.86]	$F_1, F_4$
$F_2, F_3$	[0.04 0.88 0.85 0.07]	$F_2, F_3$
$F_2, F_4$	[0.08 0.81 0.05 0.79]	$F_2, F_4$
$F_3, F_4$	[0.04 0.00 0.80 0.83]	$F_3, F_4$
$F_1, F_2, F_3$	[0.86 0.88 0.82 0.02]	$F_1, F_2, F_3$
$F_1, F_2, F_4$	[0.86 0.83 0.06 0.79]	$F_1, F_2, F_4$
$F_2, F_3, F_4$	[0.03 0.82 0.78 0.81]	$F_2, F_3, F_4$
$F_1, F_3, F_4$	[0.80 0.00 0.85 0.84]	$F_1, F_3, F_4$
$F_1, F_2, F_3, F_4$	[0.83 0.79 0.82 0.78]	$F_1, F_2, F_3, F_4$

Table 11: Results of 2<sup>nd</sup> group fault detection by fuzzy-GA when  $T_c$  is added to the measured variables.

Faults	Certainty factors	Final results
Normal	[0.00 0.00 0.00 0.00]	Normal
F <sub>1</sub>	[0.99 0.00 0.00 0.00]	F <sub>1</sub>
F <sub>2</sub>	[0.00 0.97 0.00 0.00]	F <sub>2</sub>
F <sub>3</sub>	[0.01 0.00 0.96 0.01]	F <sub>3</sub>
F <sub>4</sub>	[0.01 0.00 0.02 0.95]	F <sub>4</sub>
F <sub>1</sub> ,F <sub>2</sub>	[0.96 0.97 0.00 0.01]	F <sub>1</sub> ,F <sub>2</sub>
F <sub>1</sub> ,F <sub>3</sub>	[0.99 0.00 0.97 0.00]	F <sub>1</sub> ,F <sub>3</sub>
F <sub>1</sub> ,F <sub>4</sub>	[0.94 0.00 0.00 0.98]	F <sub>1</sub> ,F <sub>4</sub>
F <sub>2</sub> ,F <sub>3</sub>	[0.01 0.96 0.97 0.01]	F <sub>2</sub> ,F <sub>3</sub>
F <sub>2</sub> ,F <sub>4</sub>	[0.02 0.98 0.00 0.97]	F <sub>2</sub> ,F <sub>4</sub>
F <sub>3</sub> ,F <sub>4</sub>	[0.01 0.00 0.97 0.99]	F <sub>3</sub> ,F <sub>4</sub>
F <sub>1</sub> ,F <sub>2</sub> ,F <sub>3</sub>	[0.97 0.94 0.95 0.01]	F <sub>1</sub> ,F <sub>2</sub> ,F <sub>3</sub>
F <sub>1</sub> ,F <sub>2</sub> ,F <sub>4</sub>	[0.95 0.98 0.02 0.92]	F <sub>1</sub> ,F <sub>2</sub> ,F <sub>4</sub>
F <sub>2</sub> ,F <sub>3</sub> ,F <sub>4</sub>	[0.01 0.96 0.98 0.95]	F <sub>2</sub> ,F <sub>3</sub> ,F <sub>4</sub>
F <sub>1</sub> ,F <sub>3</sub> ,F <sub>4</sub>	[0.96 0.00 0.97 0.98]	F <sub>1</sub> ,F <sub>3</sub> ,F <sub>4</sub>
F <sub>1</sub> ,F <sub>2</sub> ,F <sub>3</sub> ,F <sub>4</sub>	[0.97 0.99 0.94 0.97]	F <sub>1</sub> ,F <sub>2</sub> ,F <sub>3</sub> ,F <sub>4</sub>

diagnostic systems, deviations in opposite directions from the normal condition are applied to system variables. Simulation results indicate that all faults have been detected correctly. The main advantage of the proposed scheme is the ability of detecting multiple faults. Comparison of two fuzzy systems (conventional fuzzy system and the genetic fuzzy system) reveals that both systems can diagnose the system faults, but the genetic fuzzy system diagnoses the faults with a higher certainty factor.

#### Notation

$A_T$	Heat transfer area, m <sup>2</sup>
$A_1, A_2$	Preexponential factors in Arrhenius equation
$C$	The parameter of Gaussian membership function
$C_{heat,ag}$	Heat capacity of cooling agent, J g <sup>-1</sup> K <sup>-1</sup>
$C_{heat,r}$	Heat capacity of mass of reaction, J g <sup>-1</sup> K <sup>-1</sup>
$C_{Hm}$	Concentration of inlet medium protons, mol/L
$c_j$	Concentration of ion j, j = Na, Ca, Mg, Cl, CO <sub>3</sub> , etc.
$c_{O_2}$	Oxygen concentration in the liquid phase, mg/L
$c_{O_2}^*$	Equilibrium concentration of oxygen in the liquid phase, mg/L

$c_{O_2,0}^*$	Equilibrium concentration of oxygen in distilled water, mg/L
$C_P$	Product (ethanol) concentration, g/L
$c_S$	Substrate (glucose) concentration, g/L
$c_{S,in}$	Glucose concentration in the feed flow, g/L
$c_X$	Biomass (yeast) concentration, g/L
$E_{a1}, E_{a2}$	Apparent activation energy for the growth, respectively, denaturation reaction
$F_b$	Flow of base entering the reactor, L h <sup>-1</sup>
$F_C$	Flow of cooling agent, L h <sup>-1</sup>
$F_e$	Outlet flow from the reactor, L h <sup>-1</sup>
$F_i$	Flow of substrate entering the reactor, L h <sup>-1</sup>
$F_j$	The j <sup>th</sup> fuzzy output variable
$H_{gen}$	The generation of hydrogen by micro organism, mol/L
$H_h$	The h <sup>th</sup> output linguistic value
$H_i$	Specific ionic constant of ion i (i = Na, Ca, Mg, Cl, CO <sub>3</sub> , etc.)
$I_i$	Ionic strength of ion i (i = Na, Ca, Mg, Cl, CO <sub>3</sub> , etc.)
$(k_1a)$	Product of mass-transfer coefficient for oxygen and gas-phase specific area, h <sup>-1</sup>

$(k_{1a})_0$	Product of mass-transfer coefficient at 20 °C for O <sup>2</sup> and gas-phase specific area, h <sup>-1</sup>
$K_a$	The acid dissociation constant, mol/L
$K_b$	The base dissociation constant, mol/L
$K_{O_2}$	Constant of oxygen consumption, g/L
$K_p$	Constant of growth inhibition by ethanol, g/L
$K_{p1}$	Constant of fermentation inhibition by Ethanol, g/L
$K_S$	Constant in the substrate term for growth, g/L
$K_{S1}$	Constant in the substrate term for ethanol production, g/L
$K_T$	Heat transfer coefficient, J h <sup>-1</sup> m <sup>-2</sup> K <sup>-1</sup>
$K_W$	The water dissociation constant, mol/L
$M$	The number of output fuzzy variable
$m_i$	Quantity of inorganic salt i (i = NaCl, CaCO <sub>3</sub> , MgCl <sub>2</sub> ), g
$M_i$	Molecular/atomic mass of salt/ion i, g/mol
$N$	The number of input fuzzy variable
$n_{AG}$	The number of adjustable variables
$n_{FS}$	The number of fuzzy sets for each variable
$n_{MV}$	The number of measured variable
$r_{O_2}$	Rate of oxygen consumption, mg L <sup>-1</sup> h <sup>-1</sup>
$R$	universal gas constant, 8.314 J mol <sup>-1</sup> K <sup>-1</sup>
$R_{SP}$	Ratio of ethanol produced per glucose consumed for fermentation
$R_{SX}$	Ratio of cell produced per glucose consumed for growth
$S_{i,j}$	The i <sup>th</sup> input linguistic value relevant to j <sup>th</sup> output
$Sym_i$	The i <sup>th</sup> fuzzy input variables
$T$	Time, h
$T_C$	Temperature of cooling agent in the jacket, °C
$T_{in}$	Temperature of the substrate flow entering to the reactor, °C
$T_{in,C}$	Temperature of cooling agent entering to the jacket, °C
$T_r$	Temperature in the reactor, °C
$V$	Volume of the mass of reaction, L
$V_j$	Volume of the jacket, L
$x_1$	The total concentration of the anion acid, mol/L
$x_{1,gen}$	The generation of the anion acid by micro organism, mol/L
$x_2$	The total concentration of the cation base, mol/L
$\bar{x}$	The deviation of the normal variable
$Y_{O_2}$	Yield factor for biomass on oxygen, mg/mg, defined as the amount of oxygen consumed per unit biomass produced

$Y_{x/H}$  Yield coefficient of biomass per proton, g/mol

### Greek symbols

$\alpha$	The total cation base concentration in the base flow, mol/L
$\Delta H_r$	Reaction heat of fermentation, kJ/mol O <sub>2</sub> consumed
$\sigma$	The parameter of Gaussian membership function
$\mu$	The degree of the membership function
$\mu_{O_2}$	Maximum specific oxygen consumption rate, h <sup>-1</sup>
$\mu_P$	Maximum specific fermentation rate, h <sup>-1</sup>
$\mu_X$	Maximum specific growth rate, h <sup>-1</sup>
$\rho_c$	Density of cooling agent, g/L
$\rho_r$	Density of the mass of reaction, g/L

Received : Sep. 14, 2009 ; Accepted : March 16, 2010

### REFERENCES

- [1] Venkatasubramanian V., Rengaswamy R., Yin K., Kavuri S.N., A review of Process Fault Detection and Diagnosis Part I: Quantitative Model Based Methods, *Computers and Chemical Engineering* **27**, p. 293 (2003).
- [2] Cui J., Wang Sh., Model-Based Online Fault Detection and Diagnosis Strategy for Centrifugal Chiller Systems, *International Journal of Thermal Sciences*, **44**, p. 986 (2005).
- [3] Sotomayor O.A..Z., Odloak D., Observer-Based Fault Diagnosis in Chemical Plants, *Chemical Engineering Journal*, **112**, p. 93 (2005).
- [4] Nelly Olivier-Maget, Gilles Hetreux, Jean Marc Le Lann, Marie Veronique Le Lann, Model-Based Fault Diagnosis for Hybrid Systems: Application on Chemical Processes, *Computers and Chemical Engineering*, **33**, p. 1617 (2009).
- [5] Venkatasubramanian V., Kavuri R.S.N., Yin K., A Review of Process Fault Detection and Diagnosis III: Process History Based Methods, *Computers and Chemical Engineering*, **27**, p. 327 (2003).
- [6] Genovesi A., Harmand J., Steyer J., A fuzzy Logic Based Diagnosis System for the On-Line Supervision of an Anaerobic Digester Pilot-Plant, *Biochemical Engineering Journal*, **3**, p. 171 (1999).
- [7] Eslamloueyan R., Shahrokhi M., Bozorgmehry R., Multiple Simultaneous Fault Diagnosis Via Hierarchical and Single Artificial Neural Networks, *Scientia Iranica*, **10**, p. 300 (2003).

- [8] He X.B., Yang Y.P., Yang Y.H., Fault Diagnosis Based on Variable-Weighted Kernel Fisher Discriminant Analysis, *Chemometrics and Intelligent Laboratory Systems*, **93**, p. 27 (2008).
- [9] Detroja K.P., Gudi R.D., Patwardhan S.C., Plant Wide Detection and Diagnosis Using Correspondence Analysis, *Control Engineering Practice*, **15**, p. 1468 (2007).
- [10] Cen Nan, Faisal Khan, M. Tariq Iqbal, Real-Time Fault Diagnosis Using Knowledge-Based Expert System, *Process safety and environmental protection*, **86**, p. 55 (2008).
- [11] Zadeh L.A., Making Computers Think Like People, *IEEE*, **8**, p. 26 (1984).
- [12] Binaghi E., Gallo I., Ghiselli C., Levrini L., Biondi K., An Integrated Fuzzy Logic and Web-Based Framework for Active Protocol Support, *International Journal of Medical Informatics*, **77**, p. 256 (2008).
- [13] Sheikhzadeh M., Trifkovic M., Rohani S., Fuzzy Logic and Rigid Control of a Seeded Semi-Batch, Anti-Solvent, Isothermal Crystallizer, *Chemical Engineering Science*, **63**, p. 991 (2008).
- [14] Ghoush M.A., Samhoury M., Al-Holy M., Herald T., Formulation and Fuzzy Modeling of Emulsion Stability and Viscosity of a Gum-Protein Emulsifier in a Model Mayonnaise System, *Journal of Food Engineering*, **84**, p. 348 (2008).
- [15] Mao Y., Xia Y., Yin Z., Sun Y., Wan Z., Fault Diagnosis Based on Fuzzy Support Vector Machine with Parameter Tuning and Feature Selection, *Chinese Journal of Chemical Engineering*, **15**, p. 233 (2007).
- [16] Boozarjomehry R.B., Masoori M., Which Method is Better for the Kinetic Modeling: Decimal Encoded or Binary Genetic Algorithm?, *Chemical Engineering Journal*, **130**, p. 29 (2007).
- [17] Nagy Z., Model Based Control of a Yeast Fermentation Bioreactor Using Optimally Designed Artificial Neural Networks, *Chemical Engineering Journal*, **127**, p. 95 (2007).
- [18] Pham H.T.B., Larsson G., Enfors S.O., Modeling of Aerobic Growth of *Saccharomyces Cerevisiae* in a pH-Auxostat, *Bioprocess Engineering*, **20**, p.537 (1999).