Bioprocess Engineering 16 (1997) 299–303 © Springer-Verlag 1997

A genetic-algorithm-based approach to optimization of bioprocesses described by fuzzy rules

P. Angelov, R. Guthke

Abstract A new approach to optimization of bioprocesses described by fuzzy rules is introduced in the paper. It is based on genetic algorithms (GA) and allows to determine optimal values or profiles of control variables and to optimize fuzzy rules (parameters of membership functions). The process can be described by linguistic variables and fuzzy rules. An algorithm and related software was developed. The approach was applied to an industrial antibiotic fermentation. The optimal profile of a physical variable of the preculture was determined which leads to an increasing output product concentration in the main culture of about 5%.

List of symbols

P	product (antibiotic) concentration in the main culture
X	physical variables of the preculture
N	number of sampling time units of the preculture
X_k	value of the variable X at the time point k, k = 0, ..., N
A_k	derivative of X at the time point k, k = 0,...,N
kr	time point k = Integer{r*0.1*N}, e.g. k4 = Integer{0.4*N}
H, M, L, S, SN	fuzzy linguistic variables *High, Medium, Low, Small* and *Small Negative*
p_i	singleton's values for P
μ_i	membership functions of the fuzzy sets
μ_A	resulting membership function of the derivatives A_k
μ_P	resulting membership function of the yield product concentration

Received: 27 June 1996

P. Angelov
Center for Biomedical Engineering,
Bulgarian Academy of Sciences, 105, Acad. G. Bonchev str.
Sofia -1113, Bulgaria

R. Guthke
Hans Knöll Institute for Natural Product Research, Beutenbergstr. 11 07745 Jena, Germany

Correspondence to: P. Angelov

This research was supported by a grant of the German Academic Exchange Service DAAD (ref. No. 324/14.03.96). It was carried out when the first author was a visiting fellow at the Hans Knöll Institute – Jena, Germany.

1 Introduction

The optimization problem of bioprocesses is very important for their efficiency, and it is mathematically challenging [4, 14, 16]. Deterministic descriptions commonly used are not generally satisfactory and neglect some possible process states [4, 5]. Bioprocesses are characterized by uncertainties, non-quantified factors (such as smell, taste, color, morphophysiological specifics, sedimentation rate etc.). Practically, for good performance of bioprocesses subjective estimations of the experienced technologist [16] are very important. There exist a number of generalizations of the optimization problem for the case when fuzzy elements exist [12, 13]. However, most of them are practically non-applicable to real problems, because of computational efforts [12].

In the last decade an alternative tool for optimization has been developed: the so called genetic algorithms (GA) [1, 6, 11], however their application to bioprocesses just starts.

Rivera and Karim [15] used a modified GA (the so called micro GA) for dynamic optimization of bioprocesses applying it to a neural-network-based model of ethanol production by the strain *Zymomonas mobilis* [16], however, the neural network was applied directly to the data. It would be more appropriate to the specifics of a bioprocess to use separate neural networks for each phase of the process and/or to use qualitative knowledge which could be represented by fuzzy linguistic variables.

A new approach named **FOGA** (**F**uzzy **O**ptimization supported by **G**enetic **A**lgorithms) which combines the advantages of fuzzy rule-based models and of GA is introduced in this paper. It is more *flexible* and *robust* than conventional approaches because the process can be described by linguistic variables and fuzzy rules and the derivatives can be unknown and because GA are robust with respect to the optimization problem [6]. It is tested with data of an industrial antibiotic fermentation and the optimal profile of a physical variable of the preculture is determined.

2 Fuzzy Optimization supported by Genetic Algorithms (FOGA)

The proposed approach uses the more convenient and in some cases only possible description of processes by fuzzy rules. The advantages of GA, that they don't require the structure of the model to be known, that they are robust over complex optimization surfaces and that they operate over a large number of points simultaneously, complement

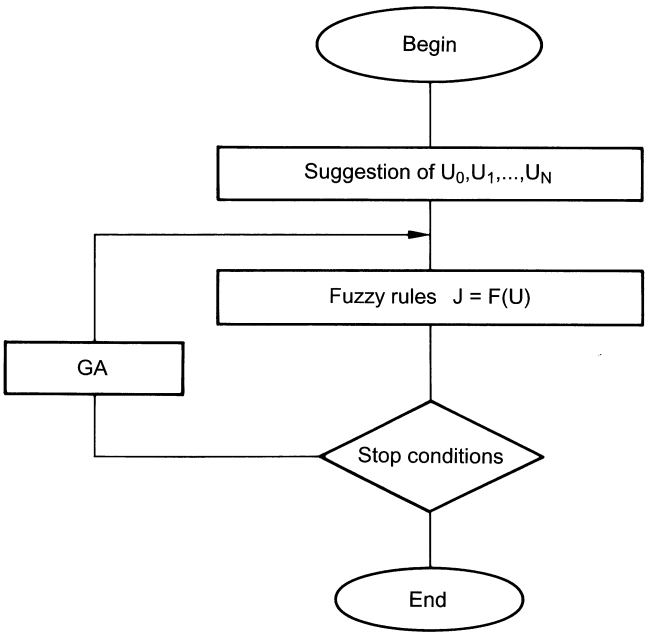

Fig. 1. Principle flow-chart of the proposed algorithm

very well the specifics of fuzzy rule-based models. On the other hand, the fuzzy models allow a more detailed description that can be closer to the nature of the processes described and can include linguistic and qualitative information. The flow-chart of the proposed approach can be sketched as shown in Fig. 1.

The algorithm is initialized by a set of chromosomes called initial population. Each chromosome consists of genes. Each gene have the value 0 or 1. The meaning of genes and respectively of the chromosomes depends on the coding procedure. Here a chromosome is a sequence of (N+1) parts each of them with **n** genes as shown in Fig. 2.

It means that the whole chromosome consist of (N+1)*n genes such that the first **n** of them are the coded value of the first unknown (U_0), the second **n** of them are the coded value of the second one (U_1), ... , and the last **n** of them are the coded value of the last one (U_N). More complex methods for coding are also possible. Each population (Π) consists of **m** chromosomes (parents) as shown in Fig. 3.

In order to minimize computation time we propose to design the initial population (Π_0) using a suggested solution (all chromosomes are initialised with the same set of suggested data) instead of random values.

At the stage of reproduction a new population (Π_1) is formed depending on the fitness value. The so called roulete wheel method [6] is used. The stage of cross over is applied at a random choosen point of the parents chromosome in order to exchange information. Mutation is a random (with probability usually less than $p_m = 0.1$) alteration of a gene value (change of 1 to 0 and vice versa). The process of generation of new populations with better fitness is repeated until a given stop-criterion is reached.

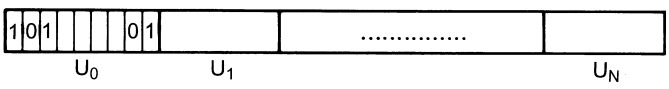

Fig. 2. Chromosome

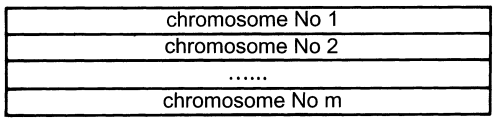

Fig. 3. Initial population

In its most general form the proposed approach can be represented by the following pseudo-computer program:

Program FOGA;
Begin
number_of_epochs: = 0
Set the initial population Π (number_of_epochs) which constitutes of values of unknowns (U_0, U_1 ,..., U_N) suggested or randomly generated into the interval of consideration;
Determine objective function's value by fuzzy rules;
while (number_of_epochs < maximal_number_of_epochs) do:
begin
number_of_epochs: = number_of_epochs +1;
Assign the probabilities to each chromosome in Π (number_of_epochs -1) which are proportional to the value of the objective function;
Generate randomly (using these probabilities) the new population Π (number_of_epochs);
Perform cross over and mutation on the genes in Π (number_of_epochs);
Calculate the objective function for each chromosome of Π (number_of_epochs);
end;
End.

3
Testing the approach with fermentation data

Experimental data of 10 fermentations of an industrial antibiotic was used in order to test the approach. The data of the same process was used in [8] for multiple correlation analysis of preculture and main fed-batch culture.

Figure 4 represents the dynamics of a physical variable (X) of the preculture (normalized values). Figure 5 shows the dynamics of the final product concentration (P) of the main culture (normalized values). The dependence of the product concentration (P) on the values of the variable X at the end of the fermentation in the preculture (X_N), on the derivative of X at the middle-time point (A_{k5}) and on the time point k8.5 ($A_{k8.5}$) are shown in Figs. 6–8 for all fermentations.

The following fuzzy rule-based model was extracted from the data:

R_1: IF (X_N is *H*) THEN (P is *H*)
R_2: IF (X_N is *L*) THEN (P is *L*)
R_3: IF (X_N is *M*) AND (A_{k5} is *L*) AND ($A_{k8.5}$ is *H*) THEN (P is *H*)
R_4: IF (X_N is *M*) AND (A_{k5} is *M*) AND ($A_{k8.5}$ is *M*) THEN (P is *M*)
R_5: IF (k1 $\leqslant$ k $\leqslant$ k4) THEN (A_k is *S*)
R_6: IF (k6 $\leqslant$ k $\leqslant$ k8) THEN (A_k is *SN*) (1)

It describes the dependance of the final product concentration at the end of the fermentation (P) on the value

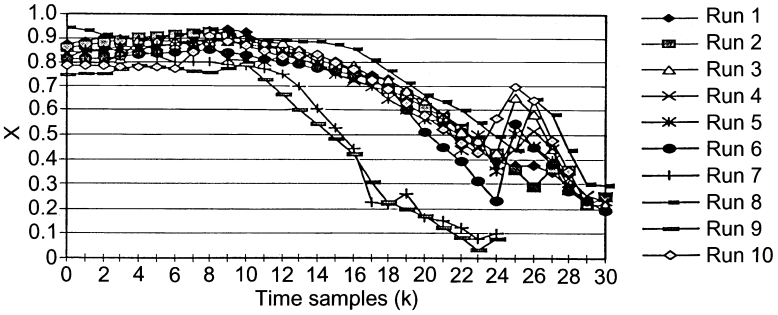

Fig. 4. Dynamics of X (N = 30)

of the physical variable (X) and its derivatives (A) in the preculture. Similar model was used in [7] for modelling the same process. We extended this model by the "dummy" rules R_5 and R_6 and modified rules R_3 and R_4. The rules R_1-R_4 describe the impact of the static parameter of the preculture (X_N) and of the derivative of X on the product concentration (P) at two time points only (at k5 and k8.5). The additional "dummy" rules R_5 and R_6 describe the dynamics of the physical variable (X) in the first phase of the process, when X is almost constant, and in the second phase, when X decreases. Such rules that describe the dynamics of the whole process are necessary in order to solve dynamic optimization problems. The experimental data are described well by this model.

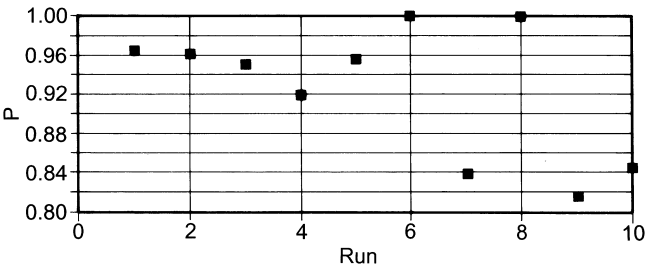

Fig. 5. Normalized values of P of all fermentation runs

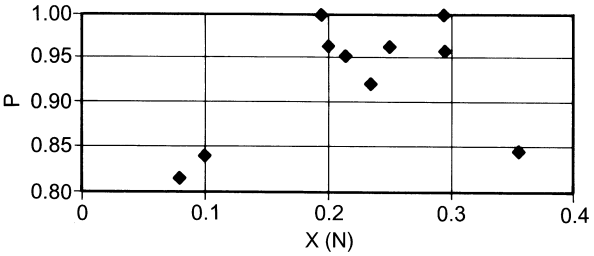

Fig. 6. Dependence of P on X_N

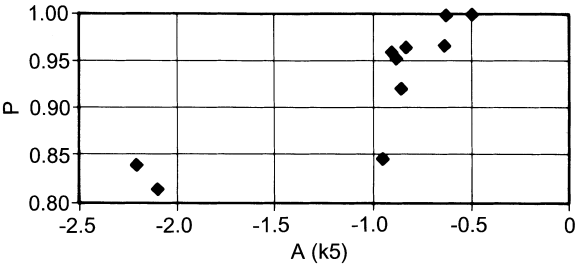

Fig. 7. Dependence of P on A_{k5}

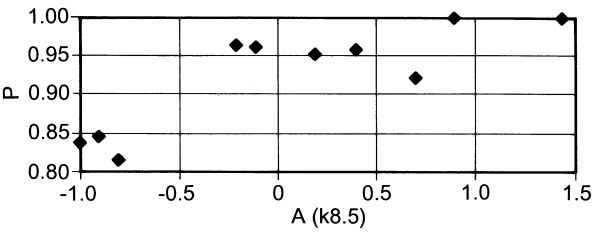

Fig. 8. Dependence of P on $A_{k8.5}$

The objective is to determine such a profile of X (X_0, X_1,..., X_N) that satisfies the highest possible output concentration of the product (P). The optimization problem can be formulated as follows:

$$J = \mu_p + \mu_A \rightarrow \max \ . \tag{2}$$

The first part of this criterion describes the resulting membership functions of the output antibiotic concentration which is determined from rules R_1-R_4. The second part describes the resulting membership functions of the derivatives of the variable X which is determined by rules R_5-R_6. It can be considered as a smooth constraint to the optimization problem whose main objective is given by the first part. The resulting membership functions can be determined by the well known operations over fuzzy sets [18] taking into account the membership functions (the degree of satisfaction) of linguistic variables in each fuzzy set, shown in Figs. 9–12, as follows:

$$\mu_p = \sum_{i=1}^{3} p_i{}^* \mu_i \tag{3}$$

$$\mu_1 = \mu_{XL}$$

$$\mu_2 = \mu_{XM}{}^* \mu_{A(k8.5)M}{}^* \mu_{A(k5)M}$$

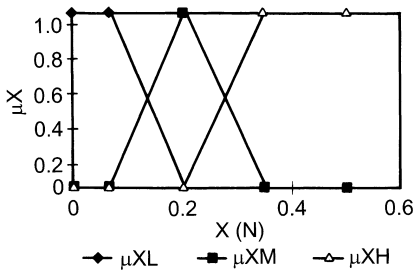

Fig. 9. Membership functions of X_N

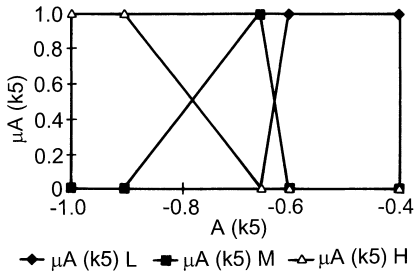

Fig. 10. Membership functions of A_{k5}

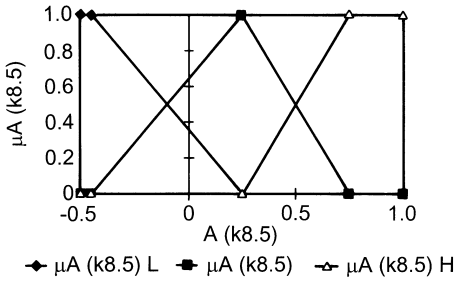

Fig. 11. Membership functions of $A_{k8.5}$

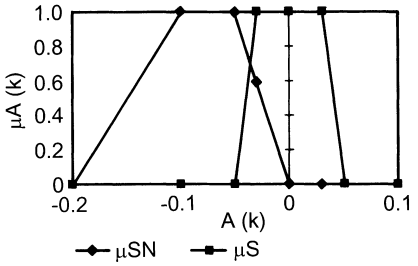

Fig. 12. Membership functions of A_k, $k1 \leqslant k \leqslant k4$ and $k6 \leqslant k \leqslant k8$

$$\mu_3 = \max[\mu_{XH}; (\mu_{XM^*}\mu_{A(k5)L^*}\mu_{A(k8.5)H})]$$

$$p = (0.804\ ;0.920\ ;1.036)^T$$

$$\mu_A = \sum_{k=kl}^{k4} \mu_S + \sum_{k=k6}^{k8} \mu_{SN} \qquad (4)$$

Applying the proposed approach we solved the problem (1)–(4) by a genetic algorithm. The initial value of X was considered to be in the interval [0.75; 0.95]. The following parameters of the genetic algorithm were used: probability for cross over (p_c = 0.3), probability for mutation (p_m = 0.05), number of chromosomes in each population (m = 40). The resulting profile of the physical variable (X) of the preculture received after 291 epochs is shown in Fig. 13. It leads to about 5% higher output product concentration than the best of the 10 experimental runs (P^{opt} = 1.05192). It can be seen that X is almost constant between time points k1 = 3 and k4 = 12 ($0.86 < X < 0.91$) and the stable decreasing of X starts after time point k4 = 12. Between the time points k6 = 18 and k8 = 24 the variable X decreases and the derivative A_k is *Small Negative* ($-0.1 < A_k < -0.05$). At the point k8.5 = 25 the derivative of X is *High* ($A_{k8.5}$ = 0.82) and at the end the optimal value of X_N is between *Medium* and *High* (X_N = 0.31). This profile is closest to one of the best experimental runs (run number 6) but at the end of the preculture the values of X are higher and this profile gives about 6% higher product concentration at the end of the main culture than in this experimental run.

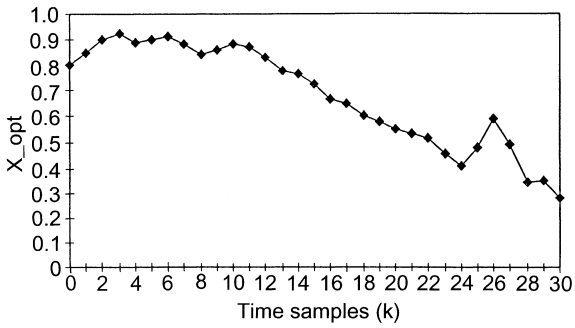

Fig. 13. Optimal profile of X

4 Conclusions

FOGA (**F**uzzy **O**ptimization supported by **G**enetic **A**lgorithms) is a new approach that combines the advantages of fuzzy rule-based models and of GA as presented in this paper. It was tested with experimental data. The optimal profile of a physical parameter of the preculture was determined which leads to increasing output product concentration of about 5%. This result demonstrates the possibility to use this approach for optimization of bioprocesses. Application of GA for adaptation of fuzzy rules is currently under investigation.

References

1. **Androulakis, I.P.; Venkatasubramanian, V.:** A genetic algorithm framework for design and optimization, Computers in Chem. Eng., 15 (1991) 217–228
2. **Angelov, P.; Tzonkov S.:** Optimal Control of Biotechnological Processes Described by Fuzzy Sets, J. Proc. Control, 3 (1993) 153–156
3. **Bettenhausen, K.; Gehlen, S.; Marenbach, P.; Tolle H.:** BioX++ - New Results and Conceptions Concerning the Intelligent Control of Biotechnological Processes. In: Proc. 6th Intern. Conf. on Comp. Applic. in Biotechnol., Garmisch-Patenkirchen 1995, pp. 324–331
4. **Breusegem, V.Van; Bastin, G.:** Optimal Control of Biomass Growth in a Mixed Culture, Biotechnol. and Bioeng., 35 (1990) 349–356
5. **Gee, D., Ramirez, E.:** Optimal Temperature Control for Batch Beer Fermentation, Biotechnol. and Bioeng., 31 (1988) 224–231
6. **Goldberg, G.E.:** Genetic Algorithms in Search, Optimization and Machine Learning. Reading, M.A., USA: Addison - Westley 1989
7. **Guthke, R.; Schmidt-Heck, W.:** Wissenserwerb für den Entwurf von Fuzzy-Regelungen bei Sekundärmetabolit-Fermentationen. In: Proc. DECHEMA Jahrestagungen'96 Wiesbaden 1996, p. 116
8. **Guthke, R.; Rausch W.:** Model Aided Multiple Correlation Analysis between Preculture and Main Fed-Batch Culture. In: Galindo, E; Ramirez O.T.: Advances in Bioprocess Engineering, pp. 267–274. The Netherlands: Kluwer Academic Publishers 1994
9. **Kacprzyk, J.:** Multistage Control of a Fuzzy System using a Genetic Algorithm. In: Proc 4th FUZZ-IEEE/IFES Congress, Yokohama 1995, pp. 1083–1088
10. **Kacprzyk, J.:** A Generalization of Fuzzy Multistage Decision Making and Control via Linguistic Quantifiers, Intl J. of Control, 38 (1983) 1249–1270

11. **Krishnakumar, K.; Goldeberg, G.E.:** Control System Optimization using Genetic Algorithms, J. of Guidance, Control and Dynamics, 15 (1990) 735–739
12. **Luhandjula, M.:** Fuzzy Optimization: An Appraisal, Fuzzy Sets and Syst. 30 (1988) 257–282
13. **Qian, Y.; Tessier, P.; Dumont, G.:** Fuzzy Logic Based Modelling and Optimization. In: Proc. 2nd Int. Conf. on Fuzzy Logic and Neural Networks, Iizuka 1992, vol.1, pp. 231–238
14. **Rivera, S.L.; Karim, M.N.:** Application of Dynamic Programming for Fermentative Ethanol Production by *Zymomonas mobilis*. In: Proc. American Control Conference, San Diego 1990, pp. 2144–2147
15. **Rivera, S.L.; Karim M.N.:** Use of Micro-Genetic Algorithms in Bioprocess Optimization. In: Proc.12th IFAC World Congress, Sydney 1993, vol.8, pp.217–220
16. **Shimizu, H.; Araki K.; Shioya S.; Suga K.:** Optimal Production of Glutahione by Controlling the Specific Growth rate in Fed-Batch Culture, Biotechnol. and Bioeng. 38 (1991) 196–205
17. **Staniskis, J.; Kildishas, V.; Filev D.:** Improvement of the Control of Biotechnological Processes. Viljnus, Lituania: LitNIINTI 1988 (in Russian)
18. **Zimmermann, H.:** Fuzzy Sets – Theory and its Application, Kluwer: Dordrecht 1985.