

Modified Multi-population Genetic Algorithm for Yeast Fed-batch Cultivation Parameter Identification

Angelova M.* , Tzonkov St., Pencheva T.

*Centre of Biomedical Engineering
Bulgarian Academy of Sciences
105 Acad. G. Bonchev Str., 1113 Sofia, Bulgaria
E-mail: maria@clbme.bas.bg*

Summary: In this work, a modified multi-population genetic algorithm is developed for the purpose of parameter identification of fermentation process model. Modified multi-population genetic algorithm is similar to the multi-population one and its development is instigated by modified genetic algorithm, similar to simple one. A comparison of four types of genetic algorithms, namely simple, modified, multi-population and modified multi-population is presented for parameter identification of a fed-batch cultivation of *Saccharomyces cerevisiae*.

Keywords: Modified multi-population genetic algorithm, Yeast fed-batch cultivation, Parameter identification.

1. INTRODUCTION

Fermentation processes are complex, nonlinear, dynamic systems with interdependent and time-varying process variables. Their modelling and high-quality control is a complicated and rather time consuming task, since their specific peculiarities lead to obtaining of non-linear models with a very complex structure. An important step for adequate modelling is the choice of a certain optimization procedure for model parameter identification. The conventional optimization methods in general can not overcome the limitations of the fermentation processes. As an alternative, genetic algorithms like stochastic global optimization method can be applied to reach a satisfactory solution for model's parameter identification.

Genetic algorithms (GA) are a direct random search technique for finding global optimal solution in complex multidimensional search space. GA are based on mechanics of natural selection and natural genetics, according Darwinian evolutionary theory [6]. They have been successfully applied to a variety of areas for solving many

* *Corresponding author*

engineering and optimization problems. GA are proved to be very suitable for the optimization of highly non-linear problems, that is why they are applied in the area of biotechnology, especially for parameter identification of fermentation process models [2, 10, 12].

Many improved variations of the standard genetic algorithm, can be found in the literature [5, 7]. One of them is the modified genetic algorithm (MGA) [11], where the selection operator is processed after performing of both crossover and mutation. In this way, the destruction of a reached good solution by either crossover, or mutation, or both, could be prevented. This modified GA applied to parameter identification of an *Escherichia coli* fed-batch fermentation model improves the optimization capability and the decision time of the algorithm. The promising results obtained there encourage the application of such a modification to the multi-population genetic algorithm (MpGA). The aim of the present investigation is the development of the modified multi-population genetic algorithm (MMpGA). MMpGA is further applied to parameter identification of a fed-batch cultivation of *S. cerevisiae*.

2. DEVELOPMENT OF MODIFIED MULTI-POPULATION GENETIC ALGORITHM

Provoked by the successful application of MGA for parameter identification of an *E. coli* fed-batch fermentation model [11], development of modified multi-population genetic algorithm (MMpGA) is the aim of this investigation. Applied modification towards MpGA is the same as applied modification in MGA towards single genetic algorithm (SGA), i.e. the reproduction is processed after both operators of crossover and mutation have covered the best solution from the current generation to be superior than or at least the same as the past.

Multi-population genetic algorithm is a single population genetic algorithm, in which many populations, called subpopulations, evolve independently from each other for a certain number of generations. After a certain number of generations (isolation time) a number of individuals are distributed between the subpopulations. There are many genetic parameters and operators [3, 4, 8], which determine the degree of genetic diversity that can occur in the subpopulations and the exchange of information between the same subpopulations. The

choice of one or other genetic parameter and operator can improve optimization solution and the decision speed.

In the beginning, the modified multi-population genetic algorithm generates a random population of n chromosomes, i.e. suitable solutions for the problem. After that individuals are reproduced through crossover and mutation. Genes from parents combine to form a whole new chromosome during the crossover and the elements of chromosome are a bit changed when a newly created offspring mutates. Then the algorithm evaluates the objective values (cost values) of the individuals in the current population. After the reproduction, the MMpGA calculates the objective function for the offspring and the best fitted individuals from the offspring are selected to replace the parents, according to their objective values. When a certain number of generations is fulfilled, a mean deviation in the population is satisfied, or when a particular point in the search space is encountered, the MMpGA is terminated.

The Matlab code for the MMpGA is listed in Fig. 1.

```
% Generational loop
while gen < MAXGEN,
% Recombine selected individuals
NewCh = recomb(XOV_F, Chrom, XOVR, SUBPOP);
% Mutate offspring
NewCh1 = mutate(MUT_F, NewCh, FieldD, MUTR, SUBPOP);
% Fitness assignment to whole population
FitnV = ranking(ObjV, [], SUBPOP);
% Select individuals from population
SelCh = select(SEL_F, NewCh1, FitnV, GGAP, SUBPOP);
% Calculate objective function for offspring
ObjVOff = feval(SelCh);
% Insert best offspring replacing worst parents
[Chrom, ObjV] = reins(Chrom, SelCh, SUBPOP, [1 INSR],
ObjV, ObjVOff);
% Increment counter
gen = gen+1;
% Migrate individuals between subpopulations
if (rem(gen, MIGGEN) == 0),
[Chrom, ObjV] = migrate(Chrom, SUBPOP, [MIGR, 1, 1],
ObjV);
end
end
```

Fig. 1 Matlab code for the MMpGA

3. PAREMETER IDENTIFICATION OF *S. CEREVISIAE* FED-BATCH CULTIVATION USING MMpGA

Experimental data of *S. cerevisiae* fed-batch cultivation is obtained in the Institute of Technical Chemistry – University of Hannover, Germany. The cultivation of the yeast *S. cerevisiae* is performed in a 2 l reactor, using a Schatzmann medium [9]. Glucose in feeding solution is 35 g/l. The temperature was controlled at 30°C, the pH at 5.5. The stirrer speed was set to 1200 rpm. The aeration rate was kept at 300 l/h. Substrate (glucose) and dissolved oxygen were measured online, while biomass and ethanol were measured offline.

The mathematical model of *S. cerevisiae* fed-batch cultivation is commonly described as follows, according to the mass balance:

$$\frac{dX}{dt} = \mu X - \frac{F}{V} X \quad (1)$$

$$\frac{dS}{dt} = -q_S X + \frac{F}{V} (S_{in} - S) \quad (2)$$

$$\frac{dE}{dt} = q_E X - \frac{F}{V} E \quad (3)$$

$$\frac{dO_2}{dt} = -q_{O_2} X + k_L^{O_2} a (O_2^* - O_2) \quad (4)$$

$$\frac{dV}{dt} = F \quad (5)$$

where X is the concentration of biomass, [g.l⁻¹]; S – concentration of substrate (glucose), [g.l⁻¹]; E – concentration of ethanol, [g.l⁻¹]; O_2 – concentration of oxygen, [%]; O_2^* – dissolved oxygen saturation concentration, [%]; F – feeding rate, [l.h⁻¹]; V – volume of bioreactor, [l]; $k_L^{O_2} a$ – volumetric oxygen transfer coefficient, [h⁻¹]; S_{in} – glucose concentration in the feeding solution, [g.l⁻¹]; μ , q_S , q_E and q_{O_2} are respectively specific rates of growth, substrate utilization, ethanol production and dissolved oxygen consumption, [h⁻¹].

Considered here fed-batch cultivation of *S. cerevisiae* is characterized with keeping glucose concentration on or below its

critical level ($S_{\text{crit}} = 0.05 \text{ g.l}^{-1}$) and with sufficient dissolved oxygen in the broth $O_2 \geq O_{2\text{crit}}$ ($O_{2\text{crit}} = 18\%$). This state corresponds to the so called mixed oxidative state, according to the functional state modeling approach [13]. As presented in [13], the specific growth rate is generally found to be a sum of two terms, one describing the contribution of sugar and the other – the contribution of ethanol to yeast growth. Both terms have the structure of Monod model. Monod model is also used for the specific ethanol and sugar consumption rates. Dissolved oxygen consumption rate is obtained as a sum of two terms, which are directly proportional to the specific glucose rate and specific ethanol production rate, respectively. That is why specific rates in equations (1)-(4) are presented as follows:

$$\mu = \mu_{2S} \frac{S}{S + k_S} + \mu_{2E} \frac{E}{E + k_E} \quad (6)$$

$$q_S = \frac{\mu_{2S}}{Y_{SX}} \frac{S}{S + k_S} \quad (7)$$

$$q_E = -\frac{\mu_{2E}}{Y_{EX}} \frac{E}{E + k_E} \quad (8)$$

$$q_{O_2} = q_E Y_{OE} + q_S Y_{OS} \quad (9)$$

where μ_{2S}, μ_{2E} – maximum growth rates of substrate and ethanol, [h^{-1}]; k_S, k_E – saturation constants of substrate and ethanol, [g.l^{-1}]; Y_{ij} – yield coefficients, [g.g^{-1}].

Mean square deviation between the model output and the experimental data obtained during cultivation has been used as an optimization criterion:

$$J_Y = \sum (Y - Y^*)^2 \rightarrow \min \quad (10)$$

where Y is the experimental data, Y^* – model predicted data, $Y = [X, S, E, O_2]$.

The presented modified multi-population genetic algorithm is applied to identification of process model parameters using *Matlab 7.0 Genetic Algorithm Toolbox*. The values of the genetic algorithm parameters are summarized in Table 1.

Table 1. Genetic parameters

Parameter	Value
GGAP	0.8
XOVR	0.95
MUTR	0.05
NIND	20
MAXGEN	100
MIGR	0.2
INSR	0.95
SUBPOP	5
MIGGEN	20

where GGAP is the generation gap – how many new individuals are created; XOVR – crossover rate; MUTR – mutation rate; PRECI – precision of binary representation; NIND – number of individuals per subpopulations; MAXGEN – maximum number of generations; MIGR – migration rate; INSR – insertion rate; SUBPOP – number of subpopulations; MIGGEN – number of generation, after which migration takes place between subpopulations.

The chosen types of genetic algorithm operators are listed in Table 2.

Table 2. Genetic operators

Operator	Type
encoding	binary
reinsertion	fitness-based
crossover	double point
mutation	bit inversion
selection	roulette wheel selection
fitness function	linear ranking

Results obtained from the parameter identification of *S. cerevisiae* fed-batch cultivation using MMpGA are conjuncted and compared to those obtained in a previous authors' investigation [1] with simple, modified and multi-population GA. Results from the comparison of four types GA are listed in Table. 3

Table 3. Results from model parameter identification

Parameter	SGA	MGA	MpGA	MMpGA
J	0.0223	0.0225	0.0147	0.0149
CPU time, s	74.7656	67.5313	146.2969	293.3750
μ_{2S} , h ⁻¹	0.9616	0.9211	0.9001	0.9001
μ_{2E} , h ⁻¹	0.0971	0.0872	0.1192	0.0600
k_S , g.l ⁻¹	0.1154	0.1176	0.1200	0.1200
k_E , g.l ⁻¹	0.7963	0.7620	0.7607	0.6506
Y_{SX} , g.g ⁻¹	0.4279	0.4279	0.4088	0.4230
Y_{EX} , g.g ⁻¹	1.2898	1.2898	6.0204	3.4794
$k_L^{O_2} a$, h ⁻¹	38.5895	127.2898	90.5778	95.4944
Y_{OS} , g.g ⁻¹	313.8285	989.8014	716.0857	750.8898
Y_{OE} , g.g ⁻¹	234.7797	62.6547	178.6444	421.4789

As it is seen from Table 3, the results obtained with four types of GA are comparable. The best value of the optimization criterion is obtained using multi-population genetic algorithms, but almost equal to obtained with modified multi-population genetic algorithm. The errors of SGA and MGA are almost equal to each other, but about 50% bigger than MpGA and MMpGA. Unfortunately, MpGA and MMpGA need more time to reach the global minimum. That is why it is up to the user to make a decision which type of GA to use as a compromise between the time consumption and model precision.

Due to the similarity of the results, the obtained results from MMpGA, developed here, will be only presented. Fig. 2-5 present results from experimental data and model simulation respectively for biomass, substrate, ethanol and dissolved oxygen.

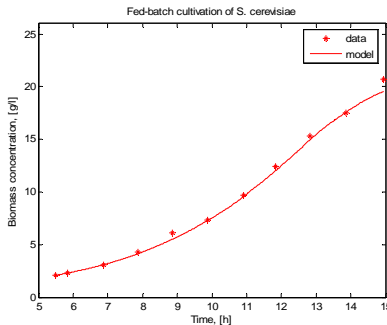


Fig. 2 Experimental and model data for biomass concentration

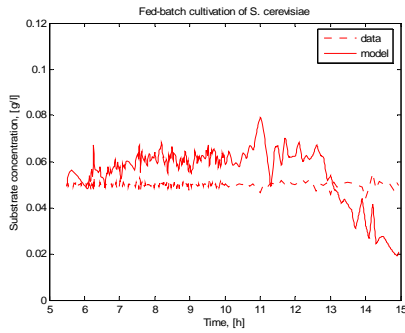


Fig. 3 Experimental and model data for substrate concentration

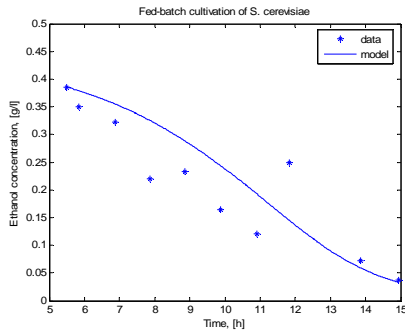


Fig. 4 Experimental and model data for ethanol concentration

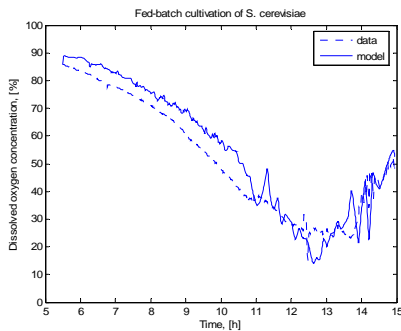


Fig. 5 Experimental and model data for dissolved oxygen concentration

The figures present the results from MMpGA application to parameter identification of *S. cerevisiae* fed-batch cultivation, which show the effectiveness of the MMpGA developed here.

4. ANALYSIS AND CONCLUSIONS

In this investigation a modified multi-population genetic algorithm has been developed and applied to parameter identification of a fed-batch cultivation of *S. cerevisiae*. The obtained results show that the introduced modification towards multi-population genetic algorithm does not lead to a significant improvement of the obtained results. The criterion value is almost equal to multi-population genetic algorithm but MMpGA are more time consuming. Expected improvement in CPU time, as it was observed with MGA towards SGA, has not been achieved. Even though the proposed modified multi-population genetic algorithm does not improve significantly the results for parameter identification of fed-batch cultivation of *S. cerevisiae*, it could be useful for some other fermentation processes.

ACKNOWLEDGEMENTS

This work is partially supported by the National Science Fund, Grant MI-1505/2005.

REFERENCES

1. Angelova M., T. Pencheva, St. Tzonkov, Parameter Identification of a Fed-batch Cultivation of *S. cerevisiae* using Genetic Algorithms, submitted in *Serdica Journal of Computing*.
2. Carrillo-Ureta G. E., P. D. Roberts, V. M. Becerra, Genetic Algorithms for Optimal Control of Beer Fermentation, *Proceeding of the 2001 IEEE International Symposium on Intelligent Control*, Mexico City, Mexico, 2001, 391–396.
3. Chipperfield A. J., P. Fleming, H. Pohlheim, C. M. Fonseca, Genetic Algorithm Toolbox for Use with MATLAB, Users's Guide, Version 1.2. *Dept. of Automatic Control and System Engineering*, University of Sheffield, UK, 1994.
4. Chipperfield A. J., P. Fleming, The Matlab Genetic Algorithm Toolbox, 1995.

5. Cordon O., F. Herrera, Hybridizing genetic algorithms with sharing scheme and evolution strategies for designing approximate fuzzy rule-based systems, *Fuzzy Sets and Systems*, 2001, 118, 235-255.
6. Goldberg D., Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wiley Publishing Company, Massachusetts, 1989.
7. Kuo R. J., C. H. Chen, Y. C. Hwang, An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network, *Fuzzy Sets and Systems*, 2001, 118, 21-45.
8. MatWorks Inc. Genetic Algorithm Toolbox, Users's Guide, 1999.
9. Pencheva T., O. Roeva, I. Hristozov, Functional state approach to fermentation processes modelling, Tzonkov St., B. Hitzmann (Eds.), Prof. Marin Drinov Academic Publishing House, Sofia, 2006.
10. Ranganath M., S. Renganathan, C. Gokulnath, Identification of Bioprocesses Using Genetic Algorithm, *Bioprocess Engineering*, 1999, 21, 123-127.
11. Roeva O., A Modified Genetic Algorithm for a Parameter Identification of Fermentation Processes, *Biotechnol. & Biotechnol.*, 2006, 20, 202-209.
12. Roeva O., Genetic Algorithm for a Parameter Estimation of a Fermentation Process Model: A Comparison, *Bioautomation*, 2005, 3, 19-28.
13. Zhang X.-C., A. Visala, A. Halme, P. Linco, Functional State Modelling Approach for Bioprocesses: Local Models for Aerobic Yeast Growth Processes, *J. Proc. Contr.*, 1994, 4(3), 127-134.