Proceedings of the 6th WSEAS Int. Conf. on EVOLUTIONARY COMPUTING, Lisbon, Portugal, June 16-18, 2005 (pp106-110)

Application of evolutionary computing for hybrid model based optimization of biochemical processes

VYTAUTAS GALVANAUSKAS, RIMVYDAS SIMUTIS, DONATAS LEVIŠAUSKAS
Process Control Department
Kaunas University of Technology
Studentų g. 50, LT-51368 Kaunas
LITHUANIA
http://www.personalas.ktu.lt/~vygal

Abstract: - General framework for hybrid model based identification and optimization of biochemical processes is elaborated. An enhanced hybrid model of the fed-batch biosurfactant production process is presented. The evolutionary programming techniques are applied for the hybrid model identification and optimization of the process. High performance of the applied optimization techniques is reported. The perspectives of development of a software package for hybrid model based identification and optimization of biochemical processes is discussed.

Key-Words: - evolutionary programming, hybrid models, model based optimization, biochemical processes

1 Introduction

Development of operational conditions for fed-batch processes in biochemical industry in order to assure high performance of these processes is of considerable importance. In the face of increased competition on the market, process optimization is a natural and straightforward choice for reducing production costs, fulfilling safety requirements, increasing process quality and reducing variability [1]. From engineering point of view such optimization includes elaboration of optimal starting conditions and profiles for control variables. Furthermore, the adjustment of the defined control policy should be made to assure the process variables fall within the range defined during the first development stage. This optimization is of particular importance in order to meet safety and operational constraints [2, 3]. Due to complexity, multi-phase and time-varying nature of the biochemical processes under consideration [1], this task may fail because of the following reasons:

1. The mechanistic and first principle models alone often can not adequately describe the process and must be enhanced. The advantages of the application of hybrid models, consisting of first principle models combined with mechanistic, ANN and fuzzy models are widely discussed [4, 5, 6, 19, 21] and their successful application in various fields of science and technology are presented [7, 8, 9].

2. Even if the first part of the problem is solved, the application of more complex models for the optimization of industrial scale processes with considerably big number of free model parameters, huge amount of data supplied for the model identification and comprehensive requirements to fulfill multiple constraints during the optimization may lead to poor performance of optimization techniques [10]. Therefore, an appropriate high performance and robust optimization technique should be implemented.

3. Optimization of the complex processes described by hybrid models proposed in [7, 8, 9] using numerical methods may lead to some specific problems which limit the range of applicable numeric methods. Discontinuities in physiological and technological constraints, complexity of the hybrid models may inhibit calculation of sensitivity functions. Therefore, the preferable optimization methods applicable for the class of the processes discussed [8, 9, 11, 12, 18] are those which do not require calculation of derivatives, e. g. simulated annealing, genetic algorithms or evolutionary programming [13, 14, 15, 20]. The application of such methods for the optimization of biochemical processes is widely discussed [12, 16].

4. In order to increase an efficiency of such model identification and optimization, general framework with user friendly environment should be developed and tested. It should include flexible description and interconnection of user blocks presented as mechanistic models, ANNs or fuzzy blocks, similar to those described in [17], powerful data management tools and robust problem oriented and easy to apply optimization routines.

The proposed approach is tackling with the bottlenecks and problems mentioned above. Nevertheless the further development of suitable software solutions is to be made.

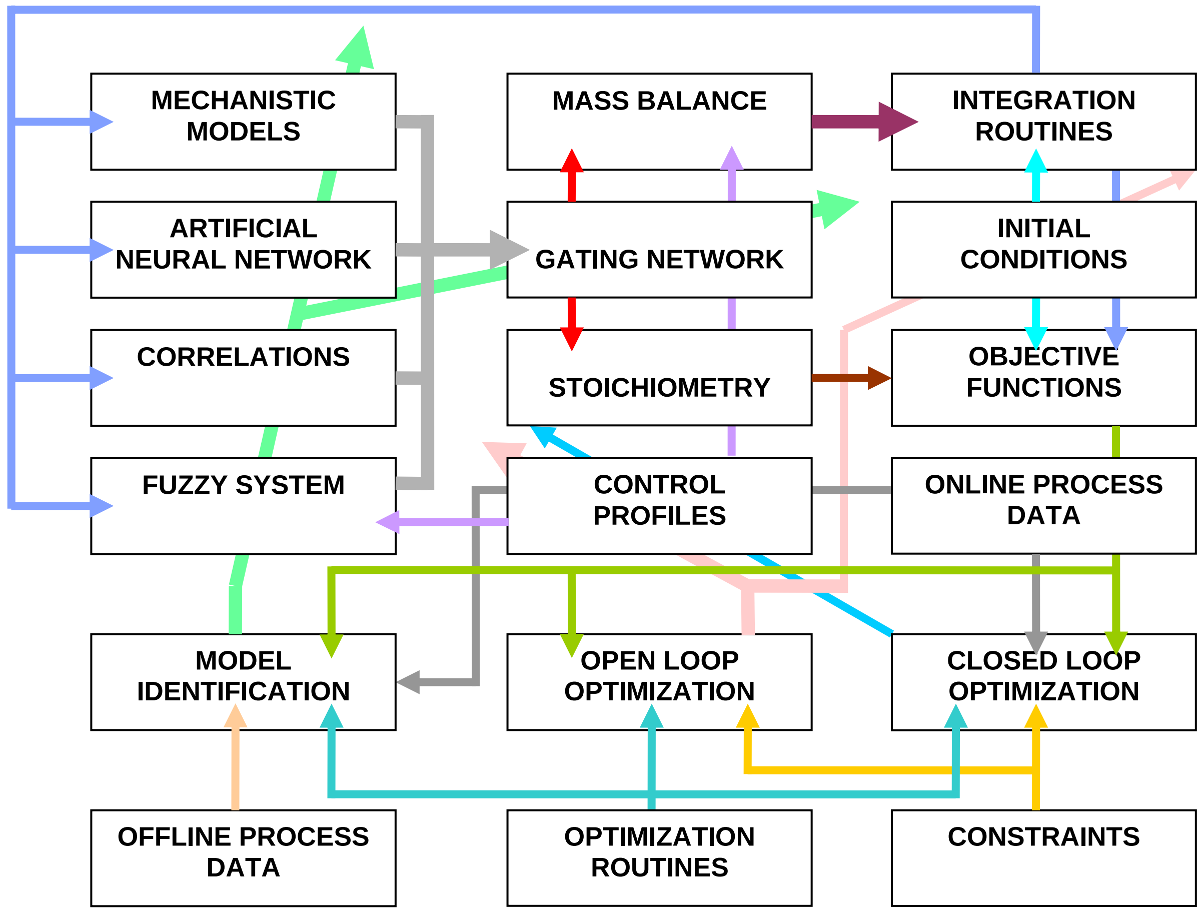

Fig. 1 General framework of the proposed hybrid model based optimization.

2 General framework of hybrid model based optimization

The proposed general framework is presented on Fig. 1 and consists of several distinct interacting parts:

a) The procedure begins with the model identification where offline process data is supplied to the optimization routines. Furthermore, at this stage the model structure should be defined, i.e. mechanistic models, engineering correlations, structure of the ANN and fuzzy blocks, mass balance equations are described by means of user friendly routines. The blocks can be interconnected between each other in parallel or sequentially and form a multilayer structure. At this stage the objective function for the procedure is defined as well. During the first raw parameter identification step only the mechanistic part of the kinetic rate expressions is tuned while the gating network is switched on and passing the output of the mechanistic part to mass balance equations.

b) Next, the input-output mapping of the state variable space should be done to ensure the application limits of the black box parts of the model. As far as the target data for the training of kinetic rates usually is not available, the outputs of the integration routine will be used to evaluate the objective function of the identification procedure. As far as neural networks are mode flexible and can better describe complex phenomena as compared to mechanistic models, the weighting of the neural network sub-model will constantly increase with the increasing amount of training data obtained after each experimental run.

c) Fuzzy subsystem is primarily intended to detect and switch between different partial models depending on the physiological state of the culture, which can be analyzed from the calculated vector of the state variables at each integration step.

d) After identification of separate parts of the kinetic model, the gating network should be trained. At this stage the weighting parameters are identified. It is important to stress that the weighting of particular parts of the model by the gating network and parameter values of the sub-models do not remain constant and are changed by the training procedure as soon as the new training data is available. After the model identification is completed, one can move to the following phase-offline optimization.

e) During the offline optimization procedure the control profiles and/or initial conditions are optimized with respect to the objective function. The latter includes functions of the state variables and stoichiometric expressions. The calculation of control profiles is done by means of standard or user defined functions, e. g. sigmoid ANNs, exponential or radial basis functions (RBF). The profiles can be explicit functions of time or divided into time grid, where switching time instants depend on the process phase or another relevant time event. The optimization is performed each time the new experimental data is added into training database in the way described in [11, 18].

f) The last stage of the framework is online optimization. The whole procedure should run online, receive the data available online, perform local retuning of the model and optimize the control actions for the next time horizon.

The framework is under development using Matlab® programming environment and standard routines as well as user defined blocks and functions.

3 Case study: fed-batch biosurfactant production process

The process of biosurfactant synthesis in fed-batch culture *Azotobacter vinelandii 21* given as an example in this paper is related to the development of a complex cleaning technology of soil contaminated by oil pollutants. More details about the bacterial strain, cultivation conditions, medium composition, experimental setup and analytical methods can be found elsewhere [8]. The aim of this optimization is to determine technological regime of the fed-batch cultivation process in order to maximize the yield of produced biosurfactant.

3.1 Hybrid model of the process

The hybrid model of the process consists of differential mass balance equations that can be written down in the following generalized form:

$$\begin{bmatrix} \frac{dC_1}{dt} \\ \dots \\ \frac{dC_i}{dt} \\ \dots \\ \frac{dC_n}{dt} \end{bmatrix} = \begin{bmatrix} q_1 \\ \dots \\ q_i \\ \dots \\ q_n \end{bmatrix} C_j + \begin{bmatrix} C_{1s}F_{1s} - C_1F \\ \dots \\ C_{is}F_{is} - C_iF \\ \dots \\ C_{ns}F_{ns} - C_nF \end{bmatrix} V^{-1} \qquad (1),$$

where C_i , q_i, C_{is}, and F_{is} is the concentration, specific reaction rate, concentration in feed and feeding rate of the *i-th* component respectively. V is culture broth volume in reactor, and F is total flow in the reactor.

The resulting values of the reaction rates $\boldsymbol{q}=f(\boldsymbol{C})$ are supplied by gating network and their current values depend on the state of the gating network and can be equal to the output of one of the sub-models or a combination of some of them.

Usually we deal with only one flow F_i corresponding to a single feeding substrate, but nevertheless it is advisable not to loose generality. Considering that for an accurate modeling it is important to account all the significant mass flows in the reactor [18], e. g. alkali or acid flow, evaporation, sampling etc., we distinguish between total flow F and flow F_i of a particular component.

3.2 Model identification

Performance index of the model identification is formulated as follows:

$$J = \sum_{i=1}^{l}\sum_{j=1}^{m}\sum_{k=1}^{n} w_{(i,j,k)}(C_{(i,j,k)} - C_{\exp(i,j,k)})^2 \rightarrow \min \qquad (2),$$

where l, m, and n are the numbers of experimental runs, data points, and variables. The values of the weights $w_{(i,j,k)}$ depend on the collected experimental data C_{exp} and reflect the differences in physical range of these variables, importance/reliability of each experiment or data point.

In this particular case, the set of data for identification consists of 3 experiments, 6 state variables and 45 measurement points each. The model includes 15 adjustable parameters.

The evolutionary programming approach was used to determine the parameter values. The iteration number was set to 10000, and the number of populations was set to 100. After identification the model was tested on new validation data. The validation results for one of the experiments are depicted on Fig. 2.

3.3 Process optimization

Performance index for the process optimization is formulated as follows:

$$J = C_5(t)V(t) \rightarrow \max, \text{ when } t=T \qquad (3),$$

where C_5 corresponds to the biosurfactant activity per volume unit and T is the fixed process time.

Additionally, the following constraints on the control variables F_i are applied:

$$0 \le F_i \le F_{i\max} \qquad (4).$$

The evolutionary programming approach was used to determine the feeding profile, and the concentrations of two components in the feeding solution.

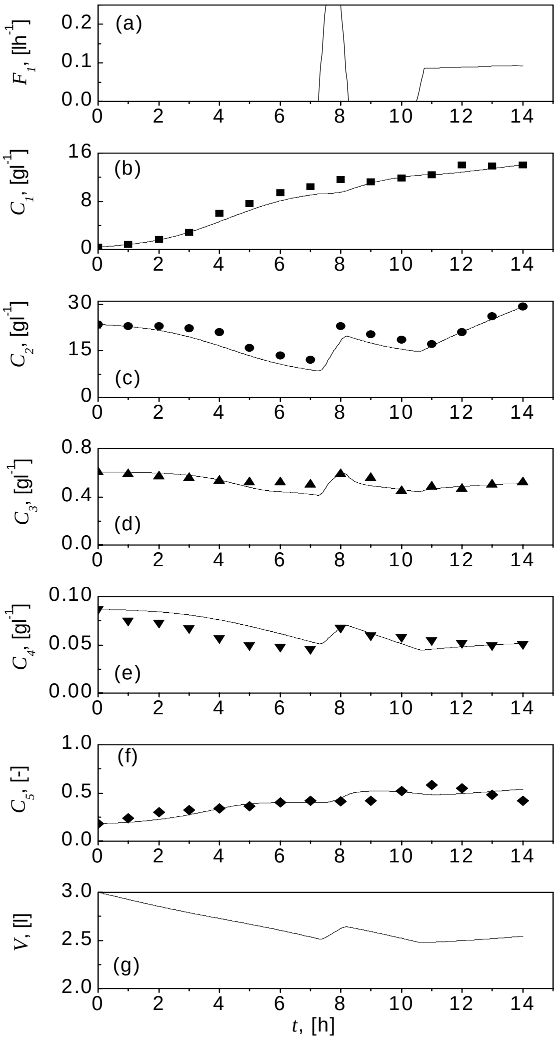

Fig. 2 Hybrid model validation results.

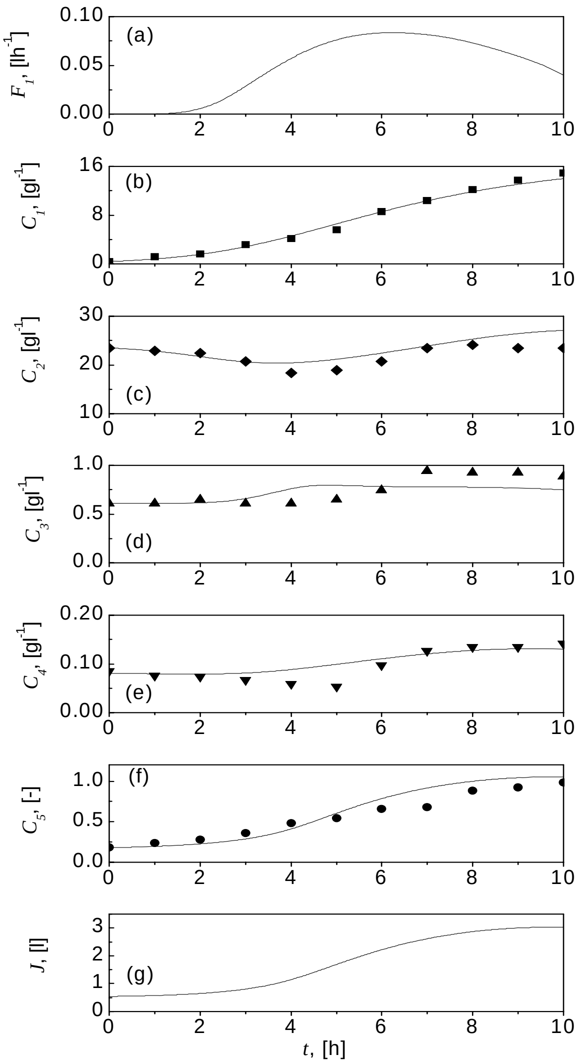

Fig. 3 Process optimization results.

The total number of 13 parameters (11 for the RBF network of the feeding rate profile, and 2 for the concentrations in feed) was used during the optimization procedure. The number of iterations was set to 2000, and the number of populations was set to 50. The optimized control strategy was validated experimentally and led to an improvement of 11% as compared to the average process performance achieved in historical data [8]. The predicted and practically achieved process optimization results are depicted on Fig. 3.

5 Conclusions

The application of the proposed approach on a laboratory scale fed-batch biochemical process showed an improvement in process performance. The optimization procedure proved to be able to tackle with strong nonlinearities of the model and presented robust behavior.

The further investigation includes software performance tests on more complex optimization tasks, incorporation of fuzzy and online optimization subsystems. Also the implementation of more optimization and integration routines would be of advantage, while dealing with more different classes of the models. The development of user friendly modeling environment is another challenging task to be solved in future.

Acknowledgement

The financial support for VG and RS in form of *NATO Reintegration Grant EAP.RIG.981479* within the *NATO Programme for Security through Science* is gratefully acknowledged.

References:

[1] D. Bonvin, Optimal operation of batch reactors - a personal view. *Journal of Process Control,* Vol.8, No.5-6, 1998, pp. 355-368.

[2] O. Abel, A. Helbig, W. Marquardt, H. Zwick, T. Daszkowski, Productivity optimization of an industrial semi-batch polymerization reactor under safety constraints, *Journal of Process Control*, Vol.10, No.4, 2000, pp. 351-362.

[3] D. Ruppen, D. Bonvin, D. W. T. Rippin, Implementation of adaptive optimal operation for a semi-batch reaction system, *Computers and Chemical Engineering*, Vol.22, 1998, pp. 185-189.

[4] A. Lübbert, R. Simutis, Adequate use of measuring data in bioprocess modeling and control, *Trends in Biotechnology*, Vol.12, 1994, pp. 304-311.

[5] S. Feyo de Azevedo, B. Dahm, F.R. Oliveira, Hybrid modelling of biochemical processes: A comparison with the conventional approach, *Computers and Chemical Engineering*, Vol.21, Suppl., 1997, pp. 751-756.

[6] V. Galvanauskas, R. Simutis, A. Lübbert, Hybrid process models for process optimisation, monitoring and control, *Bioprocess and Biosystems Engineering*, Vol.26, No.(6), 2004, pp. 393-400.

[7] S. Wermter, R. Sun, *Hybrid Neural Systems*, Springer Verlag, Heidelberg, 2000.

[8] D. Levišauskas, V. Galvanauskas, G. Žunda, S. Grigiškis, Model-based optimisation of biosurfactant production in fed-batch culture Azotobacter vinelandii, *Biotechnology Letters*, Vol.26, No.14, 2004, pp. 1141-1146.

[9] P. Georgieva, M. J. Meireles, S. Feyo de Azevedo, KBHM of fed-batch sugar crystallization when accounting for nucleation, growth and agglomeration phenomena. *Chemical Engineering Science*, Vol.58, 2003, pp. 3699-3711.

[10] Q. T. Pham, Dynamic optimization of chemical engineering processes by an evolutionary method, *Computers and Chemical Engineering*, Vol.22, 1998, pp.1089-1097.

[11] V. Galvanauskas, N. Volk, R. Simutis, A. Lübbert, Design of recombinant protein production processes, *Chemical Engineering Communications*, Vol.191, No.5, 2004, pp. 732-748.

[12] D. Levišauskas, V. Galvanauskas, R. Simutis, A. Lübbert, Model based calculation of substrate/inducer feed-rate profiles in fed-batch processes for recombinant protein production. *Biotechnology Techniques*, Vol.13, No.1, 1999, pp. 37-42.

[13] P. J. M. van Laarhoven, E. H. L. Aarts, *Simulated annealing: theory and applications*, Reidel Publishing, Dordrecht, 1987.

[14] J. H. Holland, *Adaptations in natural and artificial systems*, University of Michigan Press, Ann Arbor, 1975.

[15] D. J. Fogel, *Evolutionary computation: toward a new philosophy of machine intelligence*, IEEE Press, New York, 1995.

[16] R. Simutis, A. Lübbert, A comparative study on random search algorithms for biotechnical process optimization, *Journal of Biotechnology*, Vol.52, 1997, pp. 245-256.

[17] R. Oliveira, R. Simutis, S. Feyo de Azevedo, A. Lübbert, Hybnet, an advanced tool for process optimization and Control, *In: Proceedings of 7th Int. Conference on Computer Applications in Biotechnology – CAB7 (Eds. T. Yoshida, S. Shioya), Osaka, Japan, May 31-June 4*, 1998, pp. 315-320.

[18] V. Galvanauskas, R. Simutis, N. Volk, A. Lübbert, Model based design of a biochemical cultivation process. *Bioprocess Engineering*, Vol.18, No.3, 1998, pp. 227-234.

[19] H. A. B. Braake , J. A. Roubos, R. Babuska, Semi-mechanistic modeling and its application to biochemical processes. Chapter 10. In: H. B. Verbruggen, R. Babuska (eds.) *World Scientific Series in Robotics and Intelligent Systems*, Vol. 23, 1999. World Scientific, Signapore, 1999, pp. 205-226.

[20] J. A. Roubos, G. van Straten, A. J. B. van Boxtel, An evolutionary strategy for fed-batch bioreactor optimization; concepts and performance. *Journal of Biotechnology*, Vol.67, No.2/3, 1999, pp. 173-187.

[21] J. A. Roubos, P. Krabben, M. Setnes, R. Babuska, J. J. Heijnen, H. B. Verbruggen, Hybrid model development for fed-batch bioprocesses; combining physical equations with the metabolic network and black-box kinetics. In: H O Nyongesa, D A Linkens, M Manfouf (eds.); *6th UK Workshop on Fuzzy systems, Recent advances of fuzzy systems, nunal fuzzy, fuzzy genetic and soft computing systems*, Brunel University, Uxbridge, September 8-9, 1999. pp. 231-239.