

Multiple model approach to modelling of *Escherichia coli* fed-batch cultivation extracellular production of bacterial phytase

Olympia Roeva*

Centre of Biomedical Engineering Prof. Ivan Daskalov
Bulgarian Academy of Sciences
105 Acad. G. Bonchev Str.
Sofia 1113, Bulgaria
Tel: 359 2 979 36 48
E-mail: olympia@clbme.bas.bg

Tania Pencheva

Centre of Biomedical Engineering Prof. Ivan Daskalov
Bulgarian Academy of Sciences
105 Acad. G. Bonchev Str.
Sofia 1113, Bulgaria
Tel: 359 2 979 36 08
E-mail: tania.pencheva@clbme.bas.bg

Stoyan Tzonkov

Centre of Biomedical Engineering Prof. Ivan Daskalov
Bulgarian Academy of Sciences 105 Acad. G. Bonchev Str.
Sofia 1113, Bulgaria
Tel: 359 2 979 36 48
E-mail: tzonkov@clbme.bas.bg

Michael Arndt

Institute of Technical Chemistry
University of Hannover
LeibnizUniversity of Hannover
3 Callinstrasse, 30167
Hannover, Germany
E-mail: mail@michaelarndt.net

Bernd Hitzmann

Institute of Technical Chemistry
University of Hannover
LeibnizUniversity of Hannover
3 Callinstrasse, 30167
Hannover, Germany
E-mail: hitzmann@iftc.uni-hannover.de

Sofia Kleist

Chair of Fermentation Technique
Technical Faculty
University of Bielefeld
25 Universitätsstr., 33615
Bielefeld, Germany
E-mail: sok@fermtech.techfak.uni-bielefeld.de

Gerchard Miksch

Chair of Fermentation Technique
Technical Faculty
University of Bielefeld
25 Universitätsstr., 33615
Bielefeld, Germany
E-mail: gmi@fermtech.techfak.uni-bielefeld.de

Karl Friehs

Chair of Fermentation Technique
Technical Faculty
University of Bielefeld
25 Universitätsstr., 33615
Bielefeld, Germany
E-mail: kfr@fermtech.techfak.uni-bielefeld.de

Erwin Flaschel

Chair of Fermentation Technique
 Technical Faculty
 University of Bielefeld
 25 Universitätsstr., 33615
 Bielefeld, Germany
 E-mail: efl@fermtech.techfak.uni-bielefeld.de

Financial support: This work is partially supported from the National Science Fund Project No MI - 1505/2005.

Keywords: genetic algorithms, modelling, multiple model approach.

Abbreviations: GA: genetic algorithms

The paper presents the implementation of multiple model approach to modelling of *Escherichia coli* BL21(DE3)pPhyt109 fed-batch cultivation processes for an extracellular production of bacterial phytase. Due to the complex metabolic pathways of microorganisms, the accurate modelling of bioprocesses is rather difficult. Multiple model approach is an alternative concept which helps in modelling and control of complex processes. The main idea is the development of a model based on simple submodels for the purposes of further high quality process control. The presented simulations of *E. coli* fed-batch cultivation show how the process could be divided into different functional states and how the model parameters could be obtained easily using genetic algorithms. The obtained results and model verification demonstrate the effectiveness of the applied concept of multiple model approach and of the proposed identification scheme.

Bioprocesses, and particularly cultivation processes, are characterized by a complicated structure of organization and independent characteristics, which determine their nonlinearity and non-stationary. Model formulation for a bioprocess is traditionally performed under conditions of a well-defined medium with single-substrate limitations, conditions that are not applied to most industrial cultivations, typically running in a complex medium. In many cases, the globally valid conventional numeric models, which describe the overall process behavior cannot be used in on-line monitoring and control, either because they do not describe the process well enough or contain too many poorly known parameters. Simple unstructured models, which account for key process variables (biomass, substrate and product concentrations) do not reflect metabolic changes and are unsuitable for many tasks (Zhang et al. 1994; Tartakovsky et al. 1997; Feng and Glassey, 2000; Venkat et al. 2003). Model predictions could be improved using structured models, but these models incorporate too many equations and unknown parameters and provide a qualitative, rather than

quantitative description of the process. The structured model of a bioprocess is normally so complicated that it is difficult to be used for industrial scale production. Therefore, some alternative modelling methods for the purpose of monitoring and control of bioprocesses have to be searched for.

Table 1. Cultivation parameters.

Parameter	t_0	$\gamma_X(t_0)$	$\gamma_S(t_0)$	$V(t_0)$	γ_{S_m}	$\gamma_{S_{sp}}$
cultivation 1	4.30 h	3.20 g/l	0.78 g/l	2.70 l	500 g/l	0.2 g/l
cultivation 2	3.10 h	3.20 g/l	0.5 g/l	2.7 g/l	500 g/l	0.1 g/l

The multiple model approach is an alternative concept, which helps in modelling and control of complex processes such as bioprocesses. The state of the approaches to modelling and control problems arising working with systems of ever-increasing complexity and associated nonlinearity is presented by (Tartakovsky et al. 1997). In this work the authors describe an approach which embraces a wide range of methods by developing complex models and controllers based on multiple submodels.

The functional state concept could be use to describe and analyze the current biological state of bioprocesses, and could be applied in expert system-based fault diagnosis and in control of bioprocesses (Zhang et al. 1994). The main idea is to use a two-level hierarchy where at the first level the process is divided into macrostates, called functional states, according to behavioural equivalence. In each functional state the process is described by a conventional type of model, called a local model, which is valid only in this functional state. In each functional state certain metabolic pathways are active enough to dominate the overall behaviour of the process. The biological behaviour is quite similar during each functional state. At the second hierarchical level some numeric detection algorithms and/or rules based on expert knowledge can be used for the

*Corresponding author

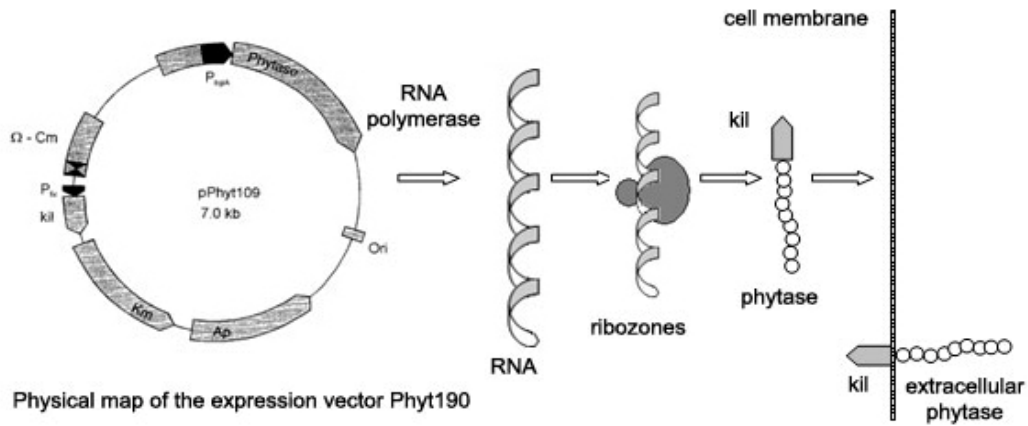


Figure 1. The expression mechanism of the intracellular phytase.

recognition of the functional states and state transitions (Zhang et al. 1994). In many batch-type processes, the functional states would be naturally identified with the different phases of the process. In a fed-batch or continuous process, the situation is more complex, but some functional states can be recognized and some functional state models can be used (Shimizu et al. 1996; Takiguchi et al. 1997).

Table 2. Rules for functional states recognition.

State	Rule
I	$\gamma_S > \gamma_{S_{crit}}$ and $w_{DO} > w_{DO_{crit}}$
II	$\gamma_S \leq \gamma_{S_{crit}}$, $w_{DO} \geq w_{DO_{crit}}$ and $\gamma_A > 0$
III	$\gamma_S \leq \gamma_{S_{crit}}$, $w_{DO} \geq w_{DO_{crit}}$ and $\gamma_A = 0$
IV	$\gamma_S < \gamma_{S_{crit}}$, $w_{DO} < w_{DO_{crit}}$ and $\gamma_A > 0$
V	$\gamma_S > \gamma_{S_{crit}}$ and $w_{DO} < w_{DO_{crit}}$

The next step after the identification of the local models is the choice of a certain optimization procedure for parameter estimation. Nowadays the most common methods used for global optimization are evolutionary algorithms such as genetic algorithms. Genetic algorithms (GA) are direct random search technique, based on the mechanics of natural selection and genetics, which can find the global optimal solution in complex multidimensional search spaces. These algorithms are proved to be very suitable for the optimization of highly non-linear problems with many variables. Recently, genetic algorithms have been extensively used for the solution of many optimization-searching problems (Goldberg, 1989; Na et al. 2002). Compared with the conventional optimization methods, GA

do not assume that the search space is differentiable or continuous. GA do not require linearity in the parameters as well which is necessary for iterative searching optimization techniques. All those properties and advantages make GA suitable for the parameter identification of cultivation processes. Many authors have reported for successful implementation of genetic algorithms in this field (Ranganath et al. 1999; Kim et al. 2002; Jeong et al. 2005; Tochamp et al. 2005). Ranganath et al. has demonstrated that GA is able to estimate the parameters of nonlinear system like fed-batch fermentor for which recursive least squares cannot be used (Ranganath et al. 1999). The authors used small population size (30 chromosomes) and therefore the algorithm is needed of 1500 iterations for evaluation. As an appropriate identification tool Roeva et al. have presented genetic algorithms used for *E. coli* MC4110 cultivation model parameters estimation (Roeva et al. 2004; Roeva, 2005). In these studies the authors have

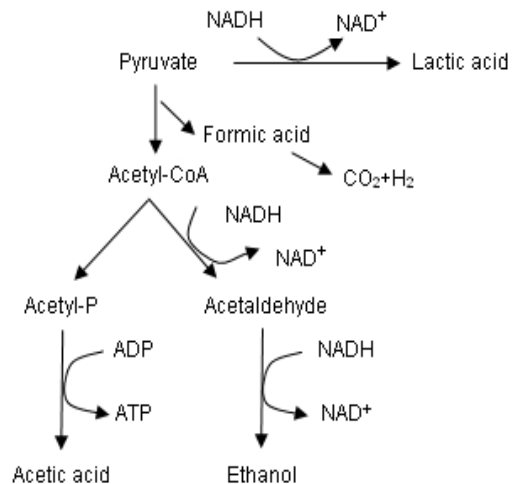


Figure 2. Fermentative metabolism of *E. coli* from glucose.

Table 3. Parameter functions of the local models in *E. coli* cultivation.

Parameter function	Functional state I	Functional state IV	Functional state V
μ	$\mu_1 \frac{\gamma_S}{k_S + \gamma_S}$	$\mu_4 \frac{\gamma_S}{k_S + \gamma_S}$	$\mu_5 \frac{\gamma_S}{k_S + \gamma_S}$
q_S	$\frac{1}{Y_{S/X}} \mu_1 \frac{\gamma_S}{k_S + \gamma_S}$	$\frac{1}{Y_{S/X}} \mu_4 \frac{\gamma_S}{k_S + \gamma_S}$	$\frac{1}{Y_{S/X}} \mu_5 \frac{\gamma_S}{k_S + \gamma_S}$
q_A	$\frac{1}{Y_{A/X}} \mu_1 \frac{\gamma_S}{k_S + \gamma_S}$	$\frac{1}{Y_{A/X}} \mu_4 \frac{\gamma_A}{k_A + \gamma_A}$	$\frac{1}{Y_{A/X}} \mu_5 \frac{\gamma_A}{k_A + \gamma_A} \frac{\gamma_S}{k_S + \gamma_S}$
q_P	$\frac{1}{Y_{P/X}} \mu_1 \frac{\gamma_P}{k_P + \gamma_P}$	$\frac{1}{Y_{P/X}} \mu_4 \frac{\gamma_S}{k_S + \gamma_S}$	$\frac{1}{Y_{P/X}} \mu_5 \frac{\gamma_S}{k_S + \gamma_S}$
q_{O_2}	$\frac{1}{Y_{DO/X}} \mu_1 \frac{w_{DO}}{k_{DO} + w_{DO}}$	$\frac{1}{Y_{DO/X}} \mu_4 \frac{w_{DO}}{k_{DO} + w_{DO}}$	$\frac{1}{Y_{DO/X}} \mu_5 \frac{w_{DO}}{k_{DO} + w_{DO}}$

used number of iterations and of individuals in a range of 100-200. The same parameters are used for models identification of methane fermentations (Kim et al. 2002; Jeong et al. 2005). In a lot of publications researchers have mainly used the following genetic parameters: crossover probability - 0.6-0.7; mutation probability - 0.02-0.05; generation gap - 0.85-0.99 (Ranganath et al. 1999; Kim et al. 2002; Jeong et al. 2005).

The main purpose of the paper is the development of a mathematical model of *E. coli BL21(DE3)pPhyt109* fed-batch cultivation for bacterial phytase extracellular production, based on the application of the multiple model approach. An estimation of the models' parameters is performed using Genetic Algorithm Toolbox procedures, especially multi-population genetic algorithm.

MATERIALS AND METHODS

The paper illustrates the concept of multiple model approach to modelling of a cultivation of *E. coli* for an extracellular production of bacterial phytase. The cultivation details (bacterial strain, culture medium and cultivation of bacteria), on-line and off-line methods, and experimental data have been published previously (Miksch et al. 2002; Kleist et al. 2003; Arndt et al. 2004). Short description of the considered here cultivation process is given below.

Fed-batch cultivation of *Escherichiacoli BL21(DE3)pPhyt109*

E. coli strain *BL21(DE3)pPhyt109* is used for cultivation experiments. The experiments are performed in the Department of Fermentation Engineering, Faculty of Technology, University of Bielefeld. Plasmid *pPhyt109*, an expression vector derived from the multi copy plasmid pUC19, contains the gene for *E. coli* phytase under the constitutive promoter of the *bglA* gene of *Bacillus amyloliquefaciens*. In addition, the expression vector contained a secretion cassette of 2.5 kb providing the competence for the secretion of pythase into the culture

```

% Generational loop
while gen < MAXGEN,
    % Fitness assignment to whole population
    FitnV = ranking(ObjV, 2, SUBPOP);
    % Select individuals from population
    SelCh = select(SEL_F, Chrom, FitnV, GGAP, SUBPOP);
    % Recombine selected individuals
    SelCh = recomb(XOV_F, SelCh, XOVR, SUBPOP);
    % Mutate offspring
    SelCh = mutate(MUT_F, SelCh, FieldD, [MUTR], SUBPOP);
    % Calculate objective function for offsprings
    ObjVOff = feval(OBJ_F, 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
    
```

Figure 3. Generational loop of a multi-population genetic algorithm.

medium based on the action of the Kil protein expressed under the control of the stationary-phase promoter of the *fic* gene (Kleist et al. 2003). The expression mechanism of the intracellular phytase is schematically presented in Figure 1.

Cultivation experiments are carried out in a bioreactor with a total volume of 7 l and a working volume of 5 l. The bioreactor is equipped with direct digital control (DDC) from MBR (Multiple Bioreactors and Sterile Plants, Zurich, Switzerland). Glucose mineral salt medium is used as growth medium. The pH is maintained at 6.9 by controlled addition of 4 N NaOH. Antifoam (PE8100, BASF, Germany) is added automatically when required. The temperature is kept at 37°C. Air flow is kept constant at 10 l min⁻¹. The stirrer speed is kept constant at 500 rpm.

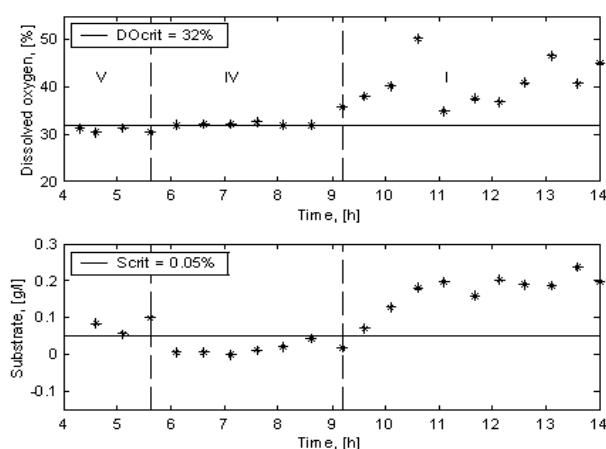


Figure 4. Recognition of functional states during *Escherichia coli* fed-batch cultivation.

Two experimental data sets are obtained and used for the modelling of *E. coli*BL21(DE3)pPhyt109 cultivation. In the first set, used for models' parameter estimation, the substrate concentration is kept at set point $\gamma_{S_{sp}} = 0.2$ g/l. The second set of experimental data, where the substrate concentration is kept at set point $\gamma_{S_{sp}} = 0.1$ g/l, is used for a model verification. The initial parameters of the considered cultivation processes are presented in Table 1.

Process modelling using multiple model approach

Modelling of bioprocesses has lead to practicing engineers being faced with problems of increasing complexity. The standard approach to complex problem solving is the applying of the divide-and-conquer strategy. A complex problem is somehow partitioned into a number of simpler sub-problems that can be solved independently and which individual solutions yield the solution of the original complex problem. The key to successful problem solution by this approach is to find suitable axes along the problem that can be partitioned. Operating regime decomposition has recently attracted significant attention. Operating

Table 4. Operators of genetic algorithm.

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

regime method leads to multiple model approaches, where different local models are applied under different operating conditions. Each model has a limited range of operating conditions in which it is sufficiently accurate or performed sufficiently well in order to serve its purpose. This range could be restricted by several factors, such as modelling assumptions, stability properties or experimental conditions. A model which is useful in a region less than the full range of operating conditions is called a local model, as opposite to a global model which is useful over the full range of operating conditions (Murray-Smith and Johansen, 1997; Petridis and Kehagias, 1998; Kordon et al. 1999).

Functional state modelling approach is originally presented (Zhang et al. 1994) for a fed-batch yeast cultivation. In general, fermentative metabolism is not the same in all microorganisms, but there are many similarities. Nielsen and Villadsen (Nielsen and Villadsen, 1994) have illustrated the analogy in the fermentative metabolism of yeasts and *E. coli*. The fermentative metabolism of *E. coli* is presented in Figure 2. Therefore, this is the precondition to implement the multiple model approach based on functional states for *E. coli* cultivation. According to the local models described by Zhang et al. 1994, this approach is applied for *E. coli* fed-batch cultivation.

The whole *E. coli* growth process can be divided into at least four functional states, according to the physiological behaviour of the microorganisms in the process (Zhang et al. 1994; Xu et al. 1999a; Xu et al. 1999b; Zelic et al. 2004).

- The first functional state (I) is the first acetate production state. The process is defined to be in

Table 5. Parameters of genetic algorithm.

Parameter	Value
generation gap	0.97
crossover rate	0.70
mutation rate	0.05
precision of binary representation	20
migration rate	0.20
insertion rate	0.90
number of subpopulations	10
number of individuals	100
number of generations	1000

this state when the sugar concentration is above the critical level ($\gamma_{S_{crit}}$) and there is sufficient dissolved oxygen. In this state acetate is produced because of high level of glucose. This process is energetically unfavorable for the cell but since the level of sugar is high it can still supply enough. The cell that is modified to produce phytase using artificial vector (*pPhyt109*) is in appropriate conditions for high phytase production. More energy activates the biosynthetic pathways in the cell and this leads to acceleration of the phytase production.

- The second functional state (II) is the mixed oxidative state. The process enters this state when the sugar concentration decreases to be equal or below the critical level and there is sufficient dissolved oxygen in the broth. Both sugar and produced acetate are cometabolised through the oxidative pathways in the state. In this state the cometabolising of the acetate and glucose leads to the utilization of more resources by the cell. The acetate metabolism process is with priority and this will lead to lower production of phytase in the cell compared to the first functional state.
- The third functional state (III) is the complete sugar oxidative state. The process is defined to be in this state when there is no acetate available, the sugar concentration is not higher than the critical level and the dissolved oxygen is above its critical level ($w_{DO_{crit}}$). In this state, sugar is completely oxidised to water and carbon dioxide. Here the complete oxidation of the substrate and the absence of side pathways activation provide

enough energy for the phytase production.

- The fourth functional state (IV) is the second acetate production state. The conditions for this state are that both concentrations, for sugar and for dissolved oxygen, to be below the corresponding critical levels. When the dissolved oxygen becomes limiting factor for microorganisms growth, acetate is produced. There are good opportunities for cell growth and activation of side anabolism pathways. Here the growth rate is higher than all other functional states and acetate levels are varying exponentially. These conditions will lead to lower phytase production.

Based on the above mentioned state description, the rules for recognition of functional states (Zhang et al. 1994) are summarized in Table 2.

Mathematical model

The following assumptions are made in the local models development of the fed-batch cultivation of *E. coli*:

- The bioreactor is completely mixed.
- Potential mixing effects of the highly concentrated feeds with the cultivation medium are neglected for the sake of the model simplicity.
- The suspension viscosity in the reactor remains constant during the experiment.
- The main products in a cultivation of *E. coli* are biomass, phytase, water, carbon dioxide and, under some conditions, acetate.
- The substrate (glucose) is consumed mainly oxidatively.
- Variations in the growth rate, acetate and phytase production, as well as in substrate consumption do not significantly change the elemental composition of biomass, thus balanced growth conditions are only assumed.
- For a simplification, the phytase production is regarded as a one-step enzymatic reaction.
- Parameters, e.g. pH and temperature, are controlled to certain acceptable constant values during the process.

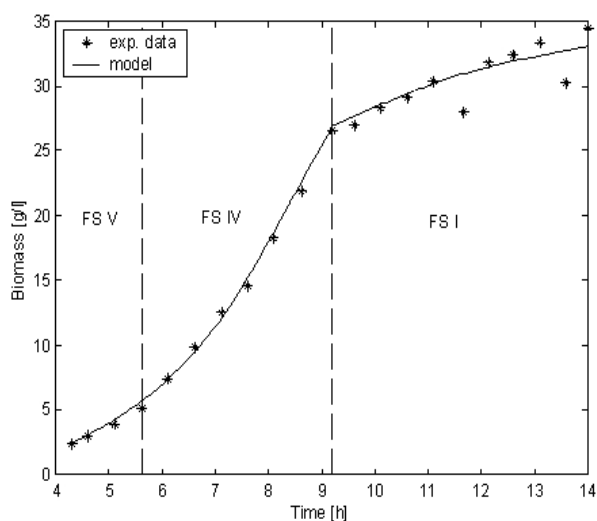


Figure 5. Simulation results and experimental data for all recognized functional states: concentration of biomass.

The rates of cell growth, sugar consumption, acetate and phytase production and dissolved oxygen concentrations in a fed-batch cultivation of *E. coli* are commonly described for all operating regimes (functional states) according to the mass balance as follows:

Table 6. Numerical values of the local model parameters.

Functional State I		Functional State IV		Functional State V	
Parameter	Value	Parameter	Value	Parameter	Value
μ_1	0.52 h ⁻¹	μ_4	0.58 h ⁻¹	μ_5	0.54 h ⁻¹
μ_A	-	μ_A	0.10 h ⁻¹	μ_A	0.14 h ⁻¹
k_S	0.076 g/l	k_S	0.006 g/l	k_S	0.04 g/l
k_P	0.10 g/l	k_A	0.51 g/l	k_A	0.10 g/l
k_{DO}	0.10%	k_{DO}	0.006%	k_{DO}	0.04%
$Y_{S/X}$	0.16 gg ⁻¹	$Y_{S/X}$	0.45 gg ⁻¹	$Y_{S/X}$	0.18 gg ⁻¹
$Y_{A/X}$	0.56 gg ⁻¹	$Y_{A/X}$	0.51 gg ⁻¹	$Y_{A/X}$	0.62 gg ⁻¹
$Y_{P/X}$	0.23 gg ⁻¹	$Y_{P/X}$	0.54 gg ⁻¹	$Y_{P/X}$	0.28 gg ⁻¹
$Y_{DO/X}$	0.45 gg ⁻¹	$Y_{DO/X}$	0.49 gg ⁻¹	$Y_{DO/X}$	0.42 gg ⁻¹
k_{La}	178.01 h ⁻¹	k_{La}	179.21 h ⁻¹	k_{La}	179.88 h ⁻¹

$$\frac{d\gamma_X}{dt} = \mu\gamma_X - \frac{Q}{V}\gamma_X \quad [1]$$

$$\frac{d\gamma_S}{dt} = -q_S\gamma_X + \frac{Q}{V}(\gamma_{S_m} - \gamma_S) \quad [2]$$

$$\frac{d\gamma_A}{dt} = q_A\gamma_X - \frac{Q}{V}\gamma_A \quad [3]$$

$$\frac{d\gamma_P}{dt} = q_P\gamma_X - \frac{Q}{V}\gamma_P \quad [4]$$

$$\frac{dw_{DO}}{dt} = -q_{DO}\gamma_X + k_{La}(w_{DO}^* - w_{DO}) - \frac{Q}{V}w_{DO} \quad [5]$$

$$\frac{dV}{dt} = Q, \quad [6]$$

where: γ_X is the concentration of biomass, g/l; γ_S - concentration of substrate (glucose), g/l; γ_A - concentration of acetate, g/l; γ_P - concentration of phytase, g/l; w_{DO} - mass fraction of dissolved oxygen, %; μ - specific growth rate, h⁻¹; q_S , q_A , q_P , q_{DO} - specific rates of substrate utilization, acetate formation, phytase formation and oxygen consumption, h⁻¹; k_{La} - volumetric oxygen transfer coefficient, h⁻¹; Q - influent flow rate, h⁻¹; V - bioreactor volume, l; γ_{S_m} - influent glucose concentration, g/l; w_{DO}^* - saturation mass fraction of dissolved oxygen, %. The

parameter functions for μ , q_S , q_A , q_P , q_{DO} in equations [1]-[5] vary in connection with the recognized functional states (Table 3).

Estimation of models parameters

For the estimation of model parameters off-line measurements of biomass, substrate (glucose), acetate, phytase, as well as on-line data for dissolved oxygen from *E. coli* fed-batch cultivation, are used. Within Matlab 5.3 environment a simulink model of the considered process has been developed. The Simulink model described by the differential equations [1]-[6], taking into account the

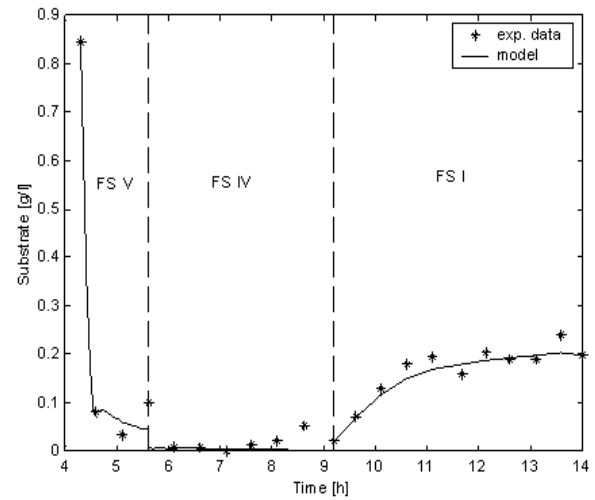


Figure 6. Simulation results and experimental data for all recognized functional states: concentration of glucose.

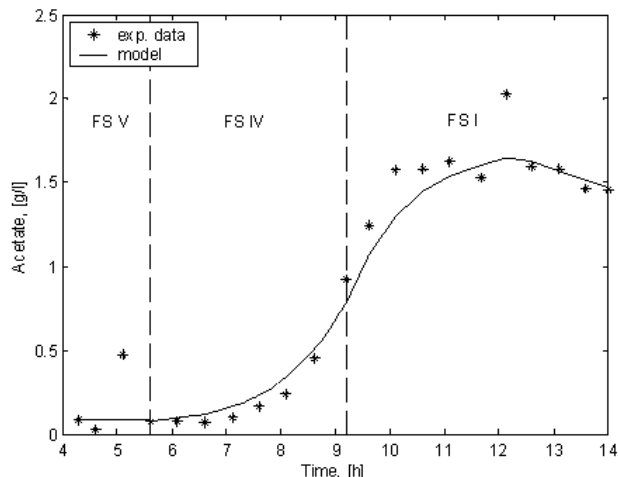


Figure 7. Simulation results and experimental data for all recognized functional states: concentration of acetate.

process parameters and the initial variable values. In order to identify the kinetic parameters in the Simulink model, a script containing necessary instructions for the genetic algorithm Toolbox has been also developed. The Matlab code for the generational loop of the multi-population GA is listed in Figure 3.

To implement the genetic algorithms, the models' parameters have to be parameterised in terms of chromosomes. Each chromosome corresponds to different objective function value. The objective function is used to provide a measure of how individuals have performed in the problem domain. In the case of minimization problem, the fitted individuals will have the lowest numerical value of the associated objective function. This raw measure of fitness is only used as an intermediate stage in determining the relative performance of individuals in genetic algorithms. The selection algorithm chooses individuals for reproduction on the basis of their relative fitness. The genetic algorithms are terminated when some criteria are satisfied. Here the termination criterion is the number of generations.

To receive better results, a single population genetic algorithm can be obtained by introducing many populations, called subpopulations (Goldberg, 1989). These subpopulations evolve independently from each other for a certain number of generations (isolation time), like the single population genetic algorithm. After the isolation time a number of individuals is distributed between the subpopulations (migration). The migration rate, the selection method of the individuals for migration and the scheme of migration determines how much genetic diversity can occur in the subpopulations and the exchange of information between subpopulations. The most general migration strategy is that of unrestricted migration (complete net topology). Here, individuals may migrate from any subpopulation to another. For each subpopulation,

a pool of potential immigrants is constructed from the other subpopulations. The individual migrants are then uniformly at random determined from this pool. The multi-population genetic algorithm models the evolution of a species in a way more similar to nature than the single population genetic algorithm.

The choice and adjustment of genetic algorithm operators and parameters were done with the criterion in order to be found the best solution with the best speed. Based on the results in (Szczerbicka et al. 1998; Roeva et al. 2004; Roeva, 2005), genetic algorithm operators and parameters for considered parameter estimation are summarized in Table 4 and Table 5.

The parameter estimation problem was stated as the minimization of a distance measure J between experimental and model predicted values of the state variables ($\gamma_x, \gamma_s, \gamma_A, \gamma_P, w_{DO}$):

$$J = \sum_{i=1}^n \sum_{j=1}^m \left\{ \left[\gamma_{exp}(i) - \gamma_{mod}(i) \right]_j \right\}^2 \rightarrow \min \quad [7]$$

where: J is the optimization criterion; $\gamma_{exp}, \gamma_{mod}$ - experimental and model data vectors; n - number of measurements for each state variable; m - number of state variables.

RESULTS AND DISCUSSION

In this work the system's full range of operation was decomposed into a number of possibly overlapping operating regimes, called functional states. In each functional state, a simple local model was applied. These local models were then combined in a way to yield a global

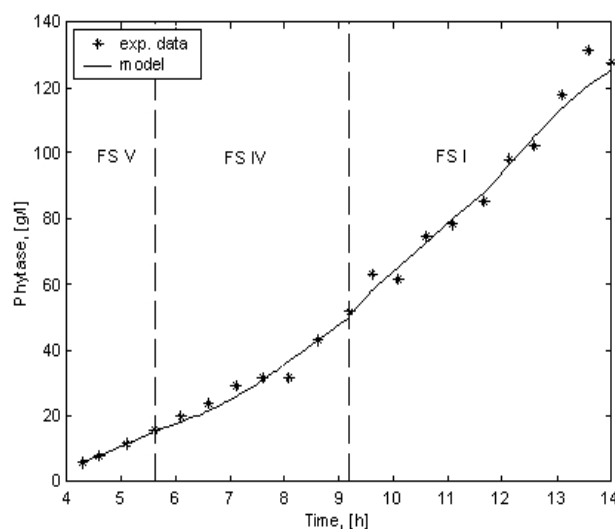


Figure 8. Simulation results and experimental data for all recognized functional states: concentration of pyruvate.

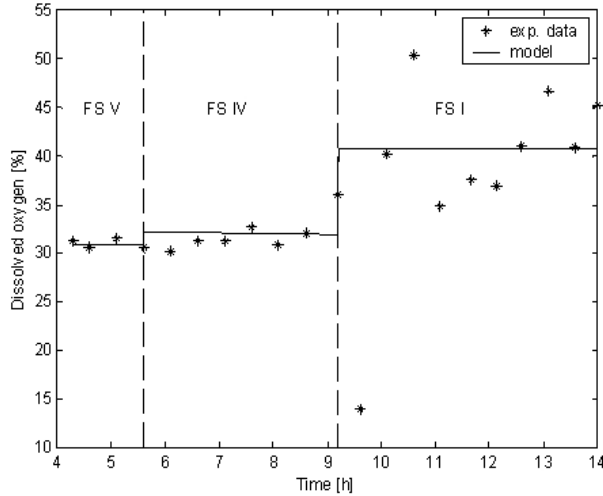


Figure 9. Simulation results and experimental data for all recognized functional states: concentration of dissolved oxygen.

model. Hence, the model development within this approach is typically consisted of the following tasks:

- Decomposition of the system's full range of operation into functional states.
- Selection of simple local model structures within each functional state. These structures could be often determined by the relevant system knowledge that is available under different operating conditions, as well as from the intended purpose of the model.
- Parameterization of the local model structures by certain variables and carrying out of parameter identification.

Based on experience (Kleist et al. 2003; Arndt et al. 2004) carried out the *E. coli* cultivations, the following values for $\gamma_{S_{crit}}$ and $w_{DO_{crit}}$ are assumed: $\gamma_{S_{crit}} = 0.05\text{g/l}$; $w_{DO_{crit}} = 32\%$. Taking into account the rules in Table 2 and the concrete values for $\gamma_{S_{crit}}$ and $w_{DO_{crit}}$ for the considered fed-batch cultivation of *E. coli*, two functional states, namely I and IV, were recognized. According to Zhang (Zhang et al. 1994) in an industrial aerobic growth process where oxygen is often limited, there might exist more functional states. For instance, a state with conditions of $w_{DO} < w_{DO_{crit}}$ and $\gamma_S > \gamma_{S_{crit}}$, and a state with $\gamma_S = 0$ and $w_{DO} < w_{DO_{crit}}$ might be possible. Although in laboratory scale cultivations those states do not occur frequently, in the considered here *E. coli* cultivation the state with conditions of $w_{DO} < w_{DO_{crit}}$ and $\gamma_S > \gamma_{S_{crit}}$

was identified. This state here was here called fifth functional state (V) - state of dissolved oxygen limitation.

The time disposition of the functional states was as follows (Figure 4). In the beginning of the cultivation the dissolved oxygen concentration was below the corresponding critical level and the glucose concentration was above its critical level. The process was in V functional state from 4.3 hrs (start of the fed-batch cultivation) to 5.6 hrs cultivation time. In the next nearly four hrs - from 5.6 hrs to 9.2 hrs cultivation time. In the next nearly four hrs - from 9.2 hrs to 14 hrs cultivation time, both the dissolved oxygen and the glucose concentrations were above the corresponding critical levels, so the process was in I functional state.

It should be noted that a fermentation process could only be in one functional state at any time. However, a certain functional state can appear in the process more than once during one run or do not appear at all. It is not obligatory all functional states, describing the physiological behaviour of the microorganisms, to appear during one cultivation. In the considered case only three functional states were identified.

At the second step, according to the specific functional state of the *E. coli* cultivation, various structures of local models were determined. Different model structures were tested. Modelling studies were performed to identify simple, easy-to-use models which to be suitable to support the engineering tasks of process optimization and control. The results from identification procedures (determination of the specific rates in local models) can be generalized as follows:

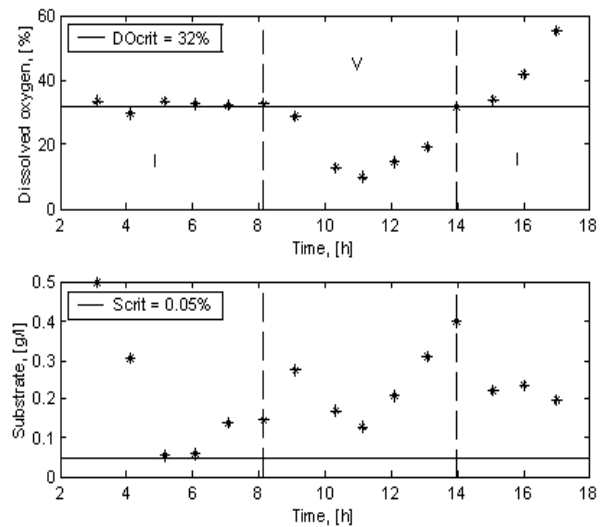


Figure 10. Recognition of functional states during *Escherichia coli* fed-batch cultivation data set used for local models verification.

- Monod kinetics was used for specific growth rate and specific sugar consumption rate in all identified functional states. The results showed that for the states IV and V the model, in which the phytase production is proportional to the specific growth rate, described in a better way the experimental data. High quality results for the functional state I were achieved by the specific phytase rate using Monod kinetics according to the phytase concentration (Table 3).
- A multiplication of two Monod kinetics was proposed for the description of acetate variation in the state V, while in the IV functional state Monod kinetics according to acetate was applied. For description of the state I the specific rate of acetate production was proposed to be proportional to the specific growth rate. Considered model structures fit quite well the dynamics of the acetate.
- In all identified functional states the variation of dissolved oxygen was described by Monod kinetics according to dissolved oxygen. The proposed model is simple and describes successfully the experimental data for dissolved oxygen concentration.

The best identification results for the local models structures of process variables were obtained for the suggestions presented in Table 3. In the presented table $\mu_1, \mu_4, \mu_5, \mu_A$ are maximum values of the specific growth rates, h^{-1} ; $\gamma_{S/X}, \gamma_{A/X}, \gamma_{P/X}, \gamma_{DO/X}$ - yield coefficients, g/g; k_S, k_A - saturation constants, g/l; k_{DO} - saturation constant, %.

Each local model contains set of definite number of model parameters that have to be defined, namely:

Functional state I $\mu_1, k_S, k_P, \gamma_{S/X}, \gamma_{A/X}, \gamma_{P/X}, \gamma_{DO/X}, k_{LA}$;
 Functional state IV $\mu_4, \mu_A, k_S, k_A, \gamma_{S/X}, \gamma_{A/X}, \gamma_{P/X}, \gamma_{DO/X}, k_{LA}$;
 Functional state V $\mu_5, \mu_A, k_S, k_A, \gamma_{S/X}, \gamma_{A/X}, \gamma_{P/X}, \gamma_{DO/X}, k_{LA}$.

Obtained results from the model parameters identification are presented in Table 6. As it could be seen, parameters values in the different functional states are different. As it is well known, the parameters of the fermentation processes models, and particularly in the cultivation of *E. coli BL21(DE3)pPhyt109*, are time-varying. The use of global process models could not reflect this fact, while the multiple model approach allows taking into account time-varying of parameters as it is shown in Table 6. In this way an adequate and more precise model is obtained.

Both the real cultivation trajectories and the simulated ones are presented in Figure 5, Figure 6, Figure 7, Figure 8 and Figure 9. Figures show the dynamics of the biomass, glucose, acetate, phytase and dissolved oxygen concentrations for all recognized functional states (I, IV and V). The initial values for the simulation in the new

functional state are the last simulated values in the previous functional state so that the trajectories have been continuous.

Obtained results clearly showed that the developed local models described the process dynamics in a high degree of accuracy. The proposed local models structures for each specific rate, relate to the corresponding functional state, fit quite well the experimental data. All assumptions about the model structures and the model parameters were confirmed. One more step to validate the developed model was to verify the model with the independent set of experimental data.

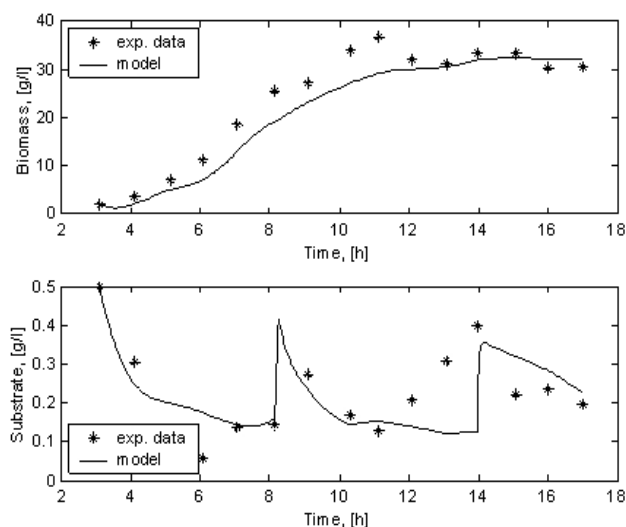


Figure 11. Simulation results from model verification for all recognized functional states: biomass and glucose.

Model verification

For the model verification was used the second data set of *E. coli BL21(DE3)pPhyt109* fed-batch cultivation with glucose set point 0.1 g/l.

The glucose and dissolved oxygen concentrations in the second data set of *E. coli* cultivation are presented in Figure 10. Based on the considered critical concentrations for glucose and dissolved oxygen and the rules listed in Table 2 three functional states were recognised. In the beginning of the cultivation, from 3.1 hrs to 8.15 hrs cultivation time was identified the functional state I, from 8.15 hrs to 13.97 hrs cultivation time - functional state V, and at the end of the process, from 13.97 hrs to 17 hrs cultivation time - again functional state I.

The developed local models for I and V functional state were tested for a prediction of the corresponding functional state behavior in the second data set. As a model input the feeding rate of the second cultivation was used and the corresponding initial conditions were taken into account.

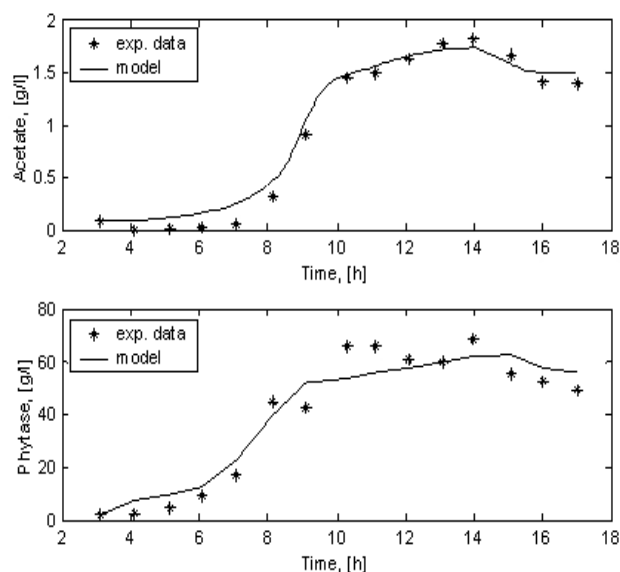


Figure 12. Simulation results from model verification for all recognized functional states: acetate and phytase.

The simulation results from the model verification are presented in Figure 11 and Figure 12.

As it can be seen the results from verification are good, with the exception of the glucose dynamics. The largest mismatch between experimental data used for verification and developed model was namely regarding to the substrate concentration. In case that the verification of rest process variables is acceptable, the failure in the glucose prediction argues for inaccurate data (noise, inexact analysis, wrong measurements, etc.). Moreover the model was developed based on data set from cultivation where the glucose was kept at value of 0.2 g/l and the data for verification are from cultivation with set point 0.1 g/l. This is also a preposition that the model could not predict the process with a high accuracy. Except unsatisfactory results about the glucose prediction, the rest of the results give a reason to conclude that the verification of the developed model was successfully fulfilled.

CONCLUDING REMARKS

Based on the application of the multiple model approach, mathematical local models of *E. coli* BL21(DE3)pPhyt109 fed-batch cultivation for a bacterial phytase extracellular production are developed. There exists a clear metabolic basis for the division into different functional states, so the characteristic dynamics in each state can be modeled with local models. To illustrate the concept of functional states, experimental data of a fed-batch cultivation of *E. coli* is used. In the considered process three functional states, namely I, IV and V are recognized. The simple structures of local models are identified for the purposes of further high quality control of *E. coli* fed-batch cultivation. Therefore, the variation of specific rates is mainly described by Monod's kinetics due to further control use. The local model

parameters are obtained using multi-population genetic algorithms. The parameter estimations show that the multiple model approach allows taking into account time-varying of parameters, while the use of global process models could not reflect this fact. The presented simulations demonstrate the effectiveness of both applied concept of the multiple model approach and proposed identification scheme. Model verification validates the developed local models and presented concept of functional states. The obtained results indicate that the multiple model approach can help in better understanding of the process behavior and simplify the process modelling.

Further investigations could be focused on modelling of non-growth associated product formation and kinetics of inhibition effects. The presented results indicate the presence of non-growth associated product formation. Namely, in functional state I the observed rate of phytase production (Figure 8) seems to be higher than the observed rate of biomass formation (Figure 5). An object for the future investigations will be modelling of product (phytase) and substrate (acetate) inhibition and development of kinetic models as model alternatives to the presented local models.

REFERENCES

- ARNDT, Michael; KLEIST, Sofia; MIKSCH, Gerhard; FRIEHS, Karl; FLASCHEL, Erwin; TRIERWEILER, J. and HITZMANN, Bernd. A feedforward-feedback substrate controller based on a Kalman filter for a fed-batch cultivation of *Escherichia coli* producing Phytase. *Computers and Chemical Engineering*, April 2004, vol. 29, no. 5, p. 1113-1120.
- FENG, M. and GLASSEY, J. Physiological state-specific models in estimation of recombinant *Escherichia coli* fermentation performance. *Biotechnology and Bioengineering*, September 2000, vol. 69, no. 5, p. 495-503.
- GOLDBERG, David E. *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley Publishing Company, Massachusetts, 1989. 413 p. ISBN 0-20-115767-5.
- JEONG, Hyeong-Seok; SUH, Chang-Won; LIM, Jae-Lim; LEE, Sang-Hyung and SHIN, Hang-Sik. Analysis and application of ADM1 for anaerobic methane production. *Bioprocess and Biosystems Engineering*, April 2005, vol. 27, no. 2, p. 81-89.
- KIM, S.; LEE, H.; KIM, J.; KO, J.; WOO, H. and KIM, S. Genetic algorithms for the application of activated sludge model no. 1. *Water Science and Technology*, 2002, vol. 45, no. 4-5, p. 405-411.
- KLEIST, Sofia; MIKSCH, Gerhard; HITZMANN, Bernd; ARNDT, Michael; FRIEHS, Karl and FLASCHEL, Erwin.

Optimization of the extracellular production of a bacterial phytase with *Escherichia coli* by using different fed-batch fermentation strategies. *Applied Microbiology and Biotechnology*, June 2003, vol. 61, no. 5-6, p. 456-462.

KORDON, Arthur; FUENTES, Yuris O.; OGUNNAIKE, Babatunde A. and DHURJATI, Prasad S. An intelligent parallel control system structure for plants with multiple operating regimes. *Journal of Process Control*, October 1999, vol. 9, no. 5, p. 453-460.

MIKSCH, Gerhard; KLEIST, Sofia; FRIEHS, Karl and FLASCHEL, Erwin. Overexpression of the phytase from *Escherichia coli* and its extracellular production in bioreactors. *Applied Microbiology and Biotechnology*, September 2002, vol. 59, no. 6, p. 685-694.

MURRAY-SMITH, Roderick and JOHANSEN, Tor A. *Multiple model approaches to modelling and control*. Taylor & Francis, London, 1997. 343 p. ISBN 0-74-840595-X.

NA, J.-G.; CHANG, Y.K.; CHUNG, B.H. and LIM, H.C. Adaptive optimization of fed-batch culture of yeast by using genetic algorithms. *Bioprocess and Biosystems Engineering*, January 2002, vol. 24, no. 5, p. 299-308.

NIELSEN, Jens and VILLADSEN, John. *Bioreaction Engineering Principles*. Plenum Press, New York and London, 1994. 540 p. ISBN 0-30-647349-6.

PETRIDIS, V. and KEHAGIAS, A. A multi-model algorithm for parameter estimation of time-varying nonlinear systems. *Automatica*, April 1998, vol. 34, no. 4, p. 469-475.

RANGANATH, M.; RENGANATHAN, S. and GOKULNATH, C. Identification of bioprocesses using genetic algorithm. *Bioprocess and Biosystems Engineering*, August 1999, vol. 21, no. 2, p. 123-127.

ROEVA, Olympia; PENCHEVA, Tania; HITZMANN, Bernd and TZONKOV, Stoyan. A genetic algorithms based approach for identification of *Escherichia coli* fed-batch fermentation. *Bioautomation*, 2004, vol. 1, p. 30-41.

ROEVA, Olympia. Genetic algorithms for a parameter estimation of a fermentation process model: a comparison. *Bioautomation*, December 2005, vol. 3, p. 19-28.

SHIMIZU, Hiroshi; MIURA, Keigo; SHIOYA, Suteaki and SUGA, Ken-ichi. On-line recognition of physiological state in a yeast fed-batch culture. *Journal of Process Control*, December 1996, vol. 6, no. 6, p. 373-378.

SZCZERBICKA, Helena; BECKER, Matthias and SYRJAKOW, Michael. Genetic algorithms: a tool for modelling, simulation, and optimization of complex

systems. *Cybernetics and Systems*, October 1998, vol. 29, no. 7, p. 639-659.

TAKIGUCHI, Noboru; SHIMIZU, Hiroshi and SHIOYA, Suteaki. An on-line physiological state recognition system for the lysine fermentation process based on a metabolic reaction model. *Biotechnology and Bioengineering*, 1997, vol. 55, no. 1, p. 170-181.

TARTAKOVSKY, B.; SHEINTUCH, M.; HILMER, J.-M. and SCHEPER, T. Modelling of *E. coli* fermentations: comparison of multicompartiment and variable structure models. *Bioprocess and Biosystems Engineering*, May 1997, vol. 16, no. 6, p. 323-329.

TOCHAMPA, Worasit; SIRISANSANEEYAKUL, Sarote; VANICHSRIRATANA, Wirat; SRINOPHAKUN, Penjit; BAKKER, Huub H.C. and CHISTI, Yusuf. A model of xylitol production by the yeast *Candida mogii*. *Bioprocess and Biosystems Engineering*, December 2005, vol. 28, no. 3, p. 175-183.

VENKAT, Aswin N.; VIJAYSAL, Prasad and GUDI, Ravindra D. Identification of complex nonlinear processes based on fuzzy decomposition of the steady state space. *Journal of Process Control*, September 2003, vol. 13, no. 6, p. 473-488.

XU, Bo; JAHIC, Mehmedalija and ENFORS, Sven-Olof. Modeling of overflow metabolism in batch and fed-batch cultures of *Escherichia coli*. *Biotechnology Progress*, February 1999a, vol. 15, no. 1, p. 81-90.

XU, B.; JAHIC, M.; BLOMSTEN, G. and ENFORS, S.-O. Glucose overflow metabolism and mixed-acid fermentation in aerobic large-scale fed-batch processes with *Escherichia coli*. *Applied Microbiology and Biotechnology*, May 1999b, vol. 51, no. 5, p. 564-571.

ZELIC, B.; VASIC-RACKI, D.; WANDREY, C. and TAKORS, R. Modeling of the pyruvate production with *Escherichia coli* in a fed-batch bioreactor. *Bioprocess and Biosystems Engineering*, July 2004, vol. 26, no. 4, p. 249-258.

ZHANG, Xia-Chang; VISALA, Arto; HALME, Aarne and LINKO, Pekka. Functional state modelling approach for bioprocesses: local models for aerobic yeast growth processes. *Journal of Process Control*, August 1994, vol. 4, no. 3, p. 127-134.