Bioautomation, 2006, 4, 12 - 23 ISSN 1312 – 451X

Modelling of Fed-batch Fermentation Process with Droppings for L-lysine Production

Tzanko Georgiev[2*], Velitchka Ivanova[1], Julia Kristeva[1], Kalin Todorov[1], Ignat Dimov[1], Alexander Ratkov[1]

[1] *Institute of Microbiology, Bulgarian Academy of Science*
Acad. G. Bonchev str., block 26, 1113 Sofia, Bulgaria, phone /fax: +359 2 870 00 97
E-mail: ratkoval@microbio.bas.bg

[2] *Technical University, Faculty of Automatics, Chair Automation of Continuous Processes*
1756 Sofia, Bulgaria, phone: +359 2 965 2408
E-mail: tzg@tu-sofia.bg

[*] *Corresponding author*

Received: December 5, 2006 **Accepted: March 15, 2006**

Published: April 26, 2006

Abstract: *The aim of the article is the development of dynamic unstructured model of L-lysine fed-batch fermentation process with droppings. This approach includes the following procedures: description of the process by generalized stoichiometric equations; preliminary data processing; identification of the specific rates (growth rate (μ), substrate utilization rate (ν), production rate (ρ)); establishment and optimization of the dynamic model of the process; simulation researches.*

Keywords: *Modelling, Optimisation, Fed-batch process with droppings, L-lysine.*

Introduction

The L-lysine is one of the important, essential amino acid. World annual production of this amino acid has been permanently increasing. Fed-batch fermentation with droppings is one of the most efficient and wildly applied types for cultivation of the microbial strain producers [4].

The synthesis of mathematical models for biotechnological processes in principal is known to be the major task of the application of modern control science for their optimisation. The models normally involve two kinds of parameters: the yield coefficients, which rely on the structure of the generalised stoichiometric reactions and the kinetic rates, which rely on the specific metabolism pathways [1].

The article aims to present the development of dynamic unstructured model for L-lysine fed-batch fermentation process with intensive droppings of the culture broth and as well as the investigation of the specificity of the process and its reflection on the obtained mathematical model.

Establishment of the dynamic unstructured model includes the following main procedures: description of the process by generalized stoichiometric equations; preliminary data processing; identification of the specific rates (growth rate (μ), substrate utilization rate (ν), production rate (ρ)); establishment and optimization of dynamic model of the process; simulation researches.

Identification procedure applied for estimation of the model structure and coefficients takes in consideration the specificity concerning dropping procedure. The important stage of this procedure is the parametric optimization of the model. The procedures for identification, optimization and simulation researches are realized by **MATLAB** and **STAGRAPHICS** packages [5, 6, 7, 8, 9]. Main approaches and steps, used for development of mathematical models are described in more details in our previous articles [2, 3].

Experimental results

Materials and methods

The fed-batch fermentation process with droppings is carried out at laboratory scale fermentor with 7 litres total volume. *Corynebacterium sp. - B031* is used as a producer. The strain is dominantly characterised with prototrophic nature, which ensures successfully carrying out of fed-batch process with big number of droppings. Analytical methods used for the characterisation of the process are as follows: biomass is measured as dry cell mass [g/l]; sugar concentration – as reducible compounds [g/l]; L-lysine – by chromatographic method. During the process on-line measurement of differed physical-chemical variables are done by proper sensors. The experimental data are shown in Fig. 1. Dissolved oxygen tension [%] is denoted as DO in Fig. 1.

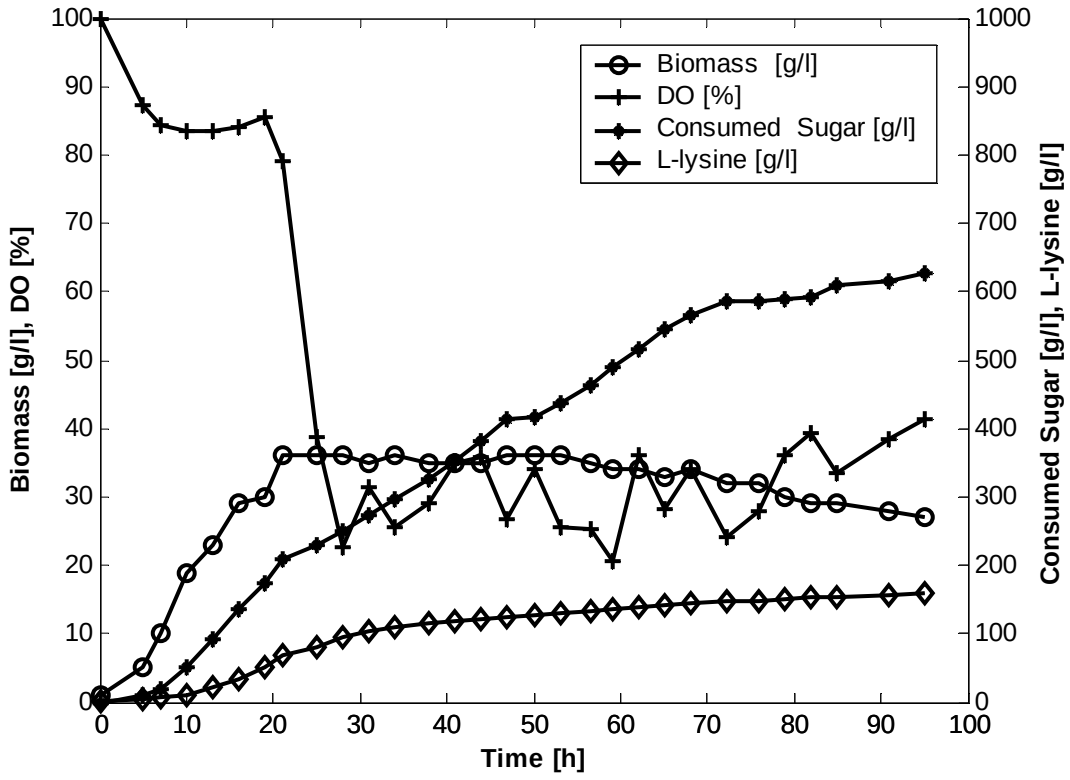

Fig. 1 Time course of the experimental data

Primary data processing

Calculation of the specific rates of the process is a final aim of this procedure. The main stages of the primary processing procedure are [2, 3]:

- Transformation of the different measurements units of the concentration to unit [g/l].
- Equalisation of the fed-batch process to the batch one.
- Calculation of the specific rates: growth rate (μ), [h^{-1}]; substrate utilisation rate, (ν) [h^{-1}]; production rate (ρ), [h^{-1}].

The specific rates are calculated by the equations:

$$\mu = \frac{\dot{X}_T}{X_T}, \quad \nu = \frac{\dot{S}_C}{X_T}, \quad \rho = \frac{\dot{L}_T}{X_T},$$

where:
$\mathbf{X_T}$ – total biomass concentration expressed by [g/l];
$\mathbf{S_C}$ – sugar consumed concentration, [g/l];
$\mathbf{L_T}$ – total L-lysine concentration, [g/l];
C – dissolved oxygen tension, [%].

The calculated specific rates are shown in Fig. 2.

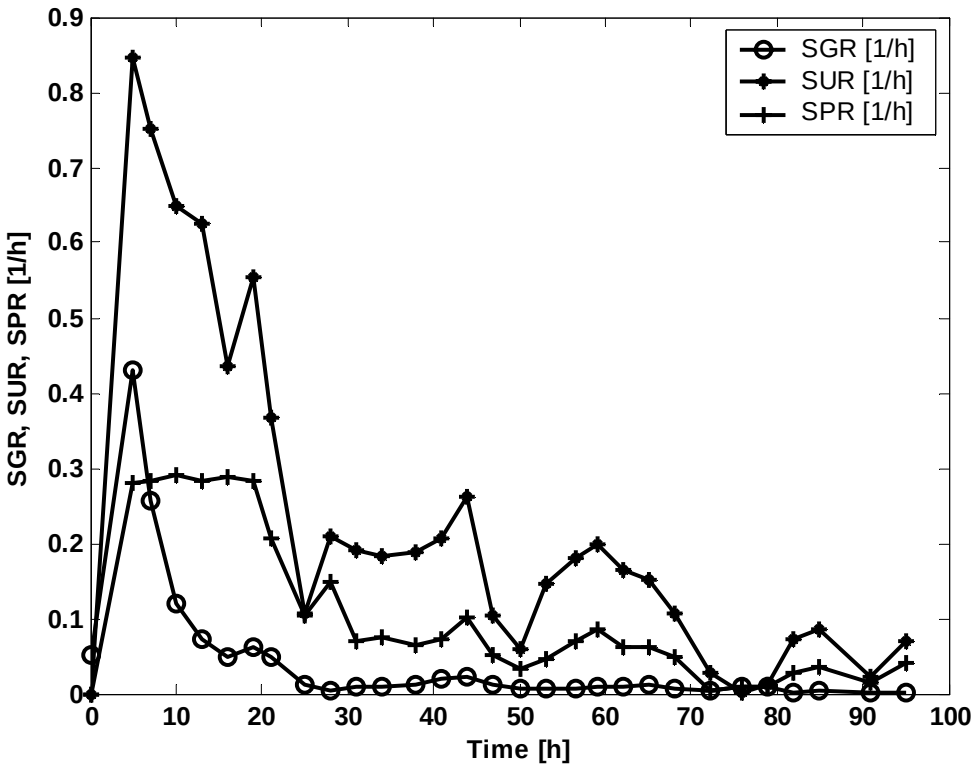

Fig. 2 The dynamic of the specific rates
SGR (instead of μ) – specific growth rate, **SUR** (instead of ν) – specific substrate utilization rate, **SPR** (instead of ρ) – specific production rate

Results and discussion

Dynamic unstructured model is the general purpose of the article. This purpose is obtained as follows.

Generalised stoichiometric equations

Generalised stoichiometric equations present a possible reactions and stages of the discussed process. These equations present a hypothesis about specific mechanisms of the product formulation.

Suppose that the fermentation process could be described by the following system of generalised stoichiometric equations:

$$\begin{aligned}
&\mathbf{S_C} \xrightarrow{\varphi_X} \mathbf{X}\\
&\mathbf{C} + \mathbf{S_C} \xrightarrow{\varphi_g} \mathbf{X} \xrightarrow{\varphi_{OUT}} \mathbf{X}\\
&\mathbf{S_R} \xrightarrow{\varphi_S} \mathbf{S_C} \xrightarrow{\varphi_{OUT}} \mathbf{S_C}\\
&\mathbf{C} + \mathbf{S_C} + \mathbf{X} \xrightarrow{\varphi_L} \mathbf{L} \xrightarrow{\varphi_{OUT}} \mathbf{L}\\
&\mathbf{V_0} \xrightarrow{\varphi_F} \mathbf{V_f} \xrightarrow{\varphi_{OUT}} \mathbf{V}
\end{aligned} \qquad (1)$$

where: $\boldsymbol{\varphi_X, \varphi_G, \varphi_S, \varphi_L, \varphi_F, \varphi_{OUT}}$ are rates of the reactions, [g/l/h];
$\mathbf{V_0}$ - initial volume, [l];
$\mathbf{V_f}$ - final volume of the culture broth, [l];
X - biomass concentration, [g/l];
S - substrate concentration as a sugar remain concentration - $\mathbf{S_R}$ or sugar consumed concentration – **Sc**, [g/l];
L - L-lysine concentration, [g/l];
C - dissolved oxygen tension, [%].
The rate $\boldsymbol{\varphi_{OUT}}$ takes into account the droppings of the culture broth.

Hypotheses about specific rates

The hypotheses concerning the specific rates of the amino acids biosynthesis are utilised as follows:

$$\begin{aligned}
&\boldsymbol{\mu = \mu(S, C)}, [h^{-1}]\\
&\boldsymbol{\nu = \nu(\mu)}, [h^{-1}]; \boldsymbol{\nu = \nu(\mu, X)}, [h^{-1}]\\
&\boldsymbol{\rho = \rho(\mu)}, [h^{-1}]; \boldsymbol{\rho = \rho(\mu, X)}, [h^{-1}].
\end{aligned} \qquad (2)$$

It is assumed that at the discrete time moments (t_k) of the dropping the derivatives of the kinetics variables are equal to zero. Semi continuous or dropping conditions are obtained based on the material balance equation as follows:

- Dropping conditions for growth

$$\mathbf{F_{OUT}(t_k) = \mu(t_k) V(t_k) - F_{IN}(t_k)}. \qquad (3)$$

- Dropping conditions for L-lysine production

$$\mathbf{F_{OUT}}(t_k) = \boldsymbol{\rho}(t_k)\frac{\mathbf{X}(t_k)}{\mathbf{L}(t_k)}\mathbf{V}(t_k) - \mathbf{F_{IN}}(t_k). \qquad (4)$$

- Dropping conditions for substrate utilization

$$\mathbf{F_{OUT}(t_k) = v(t_k)\frac{X(t_k)}{S(t_k)}V(t_k) + F_{IN}(t_k)\left(\frac{S_{IN}(t_k) - S(t_k)}{S(t_k)}\right)} \quad (5)$$

Joint conditions are obtained by comparison of the above expressions. The comparison of the equations (3) and (4) yields

$$\mathbf{\mu(t_k) = \rho(t_k)\frac{X(t_k)}{L(t_k)}} \quad (6)$$

Following the same approach the comparing of the expressions (3) and (5) obtains the equality

$$\mathbf{\mu(t_k) = v(t_k)\left(\frac{X(t_k)}{S(t_k)}\right) + \frac{F_{IN}(t_k)}{V(t_k)}\left(\frac{S_{IN}(t_k)}{S(t_k)}\right)} \quad (7)$$

The final expression is derived based on equalities (6) and (7) as follows

$$\mathbf{\rho(t_k)\frac{X(t_k)}{L(t_k)} = v(t_k)\left(\frac{X(t_k)}{S(t_k)}\right) + \frac{F_{IN}(t_k)}{V(t_k)}\left(\frac{S_{IN}(t_k)}{S(t_k)}\right)} \quad (8)$$

It could be emphasized that these conditions are satisfied at the discrete time moments (t_k).

Identification procedure

The presented approach of the identification procedure includes the following stages:

- The linear regression or polynomial regressions are applied for selection of a preliminary structure of the models describing the specific rates and initial estimates of its parameters. The aim of this step is the selection of the appropriate model structure and the model fit to the experimental data. Experimental data transformations on this step are natural logarithm and an appropriate power of the exponential terms.
- The next step is done by a non-linear regression based on the selected model structure and initial values of the parameters without any transformations. These models are represented in the article. The model selection is done based on R^2 coefficient model fit approximation and the results of the residual investigation.
- The final stage of identification is connected with the parametric optimisation of the models through the non-linear optimisation procedure under the confidence intervals of the parameters using Optimisation Toolbox. The Levenberg-Marquardt algorithm with least squares objective function is used for optimisation.

Models of the specific rates

After the identification procedure the specific growth rate is expressed as:

$$\mathbf{\mu = exp\left(a_0 + (a_1 S_C) + (a_2 S_C^2) + (b_1 C) + (b_2 C^2) + (b_3 (S_C C))\right)} \quad (9)$$

The adequacy of the model (9) graphically presented in Fig. 3 is proved through the value of the determination coefficient R^2 = 0,897692 obtained by the non-linear regression. The derived model is selected from the set of the models suitable for the experimental data subject to requirement for a minimal order of the polynomials in the model.

Table 1. Estimated parameters according to the model (9) with 95% confidence intervals

Parameters	*Estimate*	Asymptotic standard error	Lower limit	Upper limit
a_0	-10,7221	39,7702	- 92,804	71,3598
a_1	0,00116321	0,107516	- 0,220739	0,223065
a_2	0,0000174146	0,000101979	- 0,00019306	0,000227889
b_1	34,1904	92,8428	- 157,428	225,809
b_2	- 25,9271	81,0916	- 193,292	141,438
b_3	- 0,0486574	0,11946	- 0,295211	0,197896

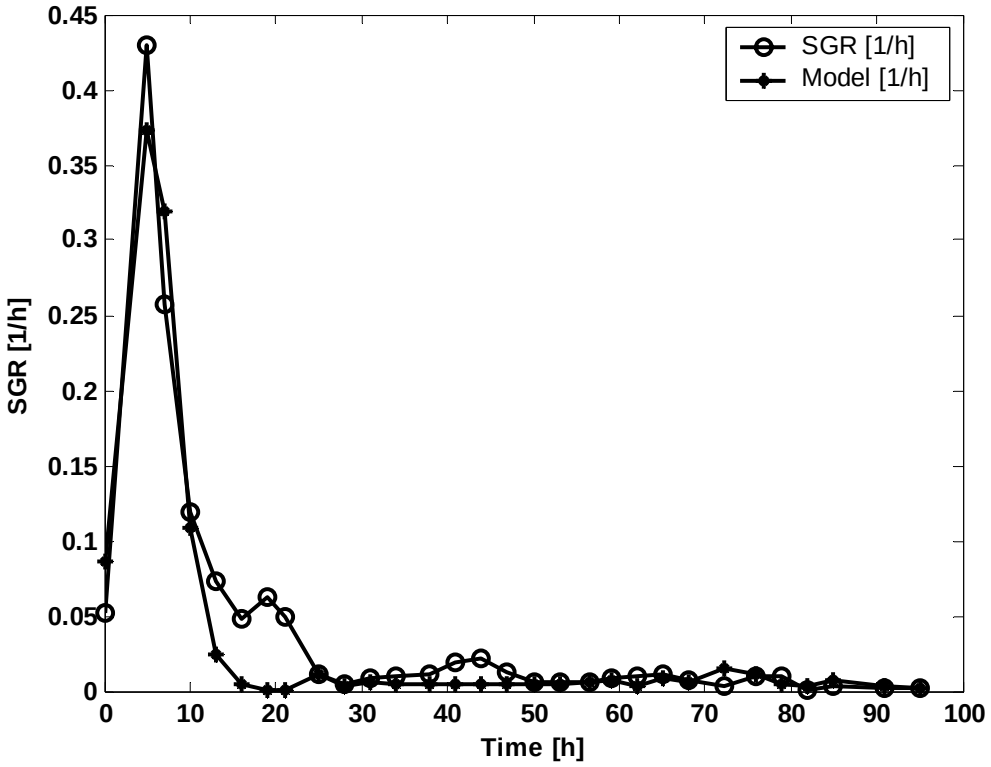

Fig. 3 Model approximation fit of the specific growth rate

Following the same approach the model of the specific utilization rate is obtained as:

$$\mathbf{v} = \exp\left(c_0 + (c_1\mu) + (c_2\mu^2) + (c_3\mu^3) + (c_4 X) + (c_5 X^2) + (c_6 X^3) + (c_7 \mu . X)\right) \quad (10)$$

The value of the determination coefficient (R^2 = 0,810537) proves the adequacy of this model (10).

The estimated parameters as a result of the non-linear regression are presented in Table 2.

Table 2. Estimated parameters according to the model (10) with 95% confidence intervals

Parameters	*Estimate*	Asymptotic standard error	Lower limit	Upper limit
c_0	- 23,7452	21,1273	- 67,5605	20,0702
c_1	51,1571	63,5027	- 80,5396	182,854
c_2	- 41,9735	265,033	- 591,62	507,673
c_3	30,9143	191,832	- 366,921	428,75
c_4	2,00615	2,06736	- 2,28131	6,29361
c_5	- 0,0648347	0,0858439	- 0,242865	0,113195
c_6	0,000724185	0,000330371	0,0000390356	0,00140933
c_7	- 0,735849	1,50274	-3,85235	2,38065

The appropriate structure of the model and the estimates of the parameters are conformed by the plots in Fig. 4.

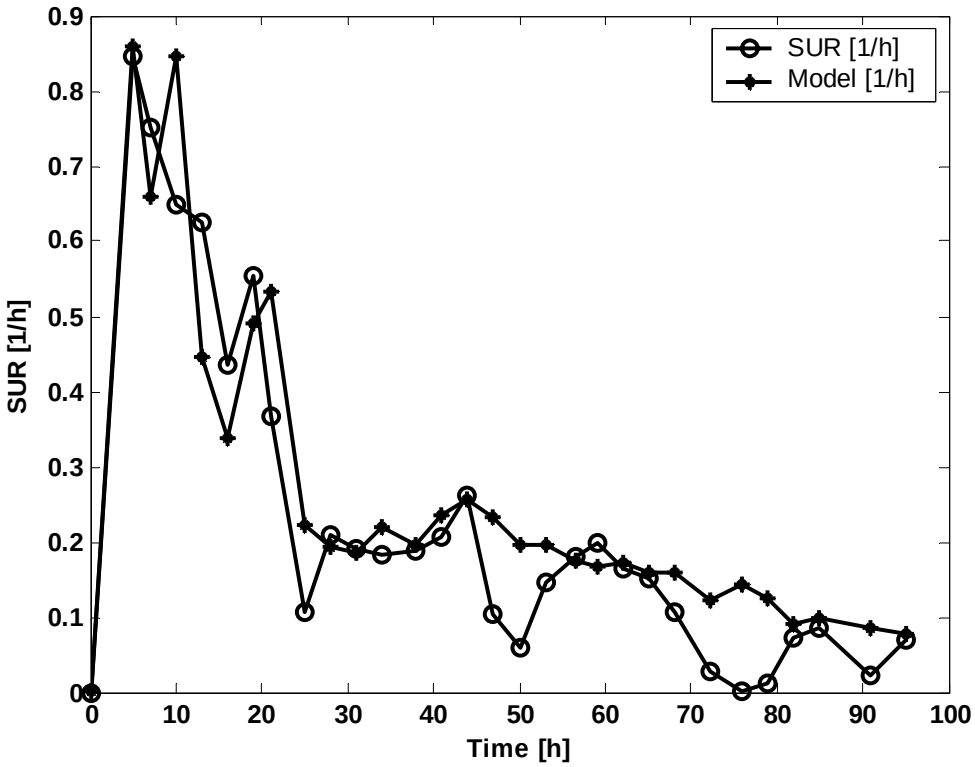

Fig. 4 Model approximation fit of the specific utilization rate

The important characterization of the L-lysine production is the specific production rate derived by the identification procedure as follows

$$\rho = \exp\left(d_0 + (d_1\mu) + (d_2\mu^2) + (d_3\mu^3) + (d_4 X) + (d_5 X^2) + (d_6 X^3) + (d_7 \mu.X)\right) \quad (11)$$

The non-linear regression yields the estimates of the parameters with confidence intervals (Table 3). The adequacy of the model is confirmed by the value of the determination coefficient $R^2 = 0{,}839314$.

Table 3. Estimated parameters according to the model (11) with 95% confidence intervals

Parameters	*Estimate*	Asymptotic standard error	Lower limit	Upper limit
d_0	- 10,5439	28,4196	- 69,4828	48,3949
d_1	51,4282	53,643	- 59,8209	162,677
d_2	-127,521	289,047	- 726,969	471,928
d_3	123,707	262,868	- 421,448	668,863
d_4	0,36768	2,56514	- 4,95211	5,68747
d_5	- 0,0017355	0,0705484	- 0,148044	0,144573
d_6	- 0,000075946	0,00033325	- 0,000767066	0,000615174
d_7	- 0,407434	1,3511	- 3,20945	2,39459

It could be seen that the model approximation fit describes the trend of the specific production rate (Fig. 5).

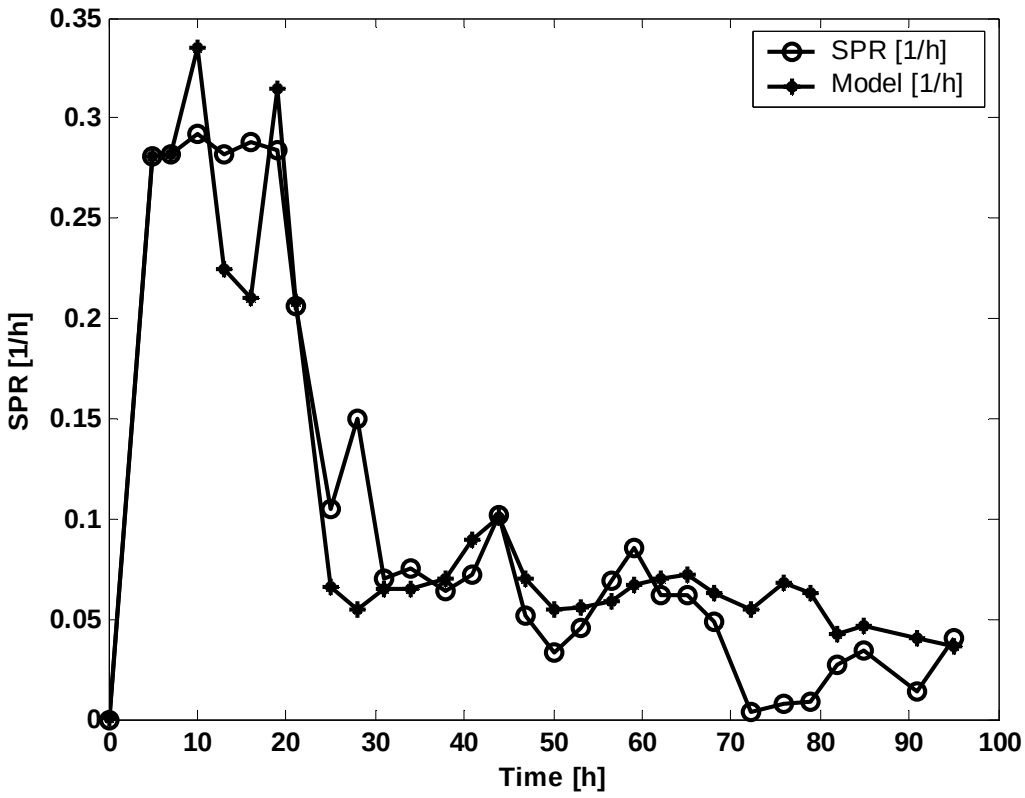

Fig. 5 Model approximation fit of the specific production rate

The obtained models with structure and estimates of the parameters are used on the next step of the synthesis of the model.

Unstructured dynamic mathematical model of the process

Base on the previous results the input–output model of the investigated process is as follows:

$$\frac{dX}{dt} = K_1 \mu X - \frac{F_{IN}}{V} X - \frac{F_{OUT}}{V} X$$
$$\frac{dS_C}{dt} = K_2 \nu X - \frac{F_{IN}}{V} S_C + \frac{F_{IN}}{V} S_{IN} - \frac{F_{OUT}}{V} S_C \quad (12)$$
$$\frac{dL}{dt} = K_3 \rho X - \frac{F_{IN}}{V} L - \frac{F_{OUT}}{V} L$$
$$\frac{dV}{dt} = F_{IN} - F_{OUT}$$

where:
X – biomass concentration, [g/l];
L – L-lysine concentration, [g/l];
Sc – sugar consumed concentration, [g/l];
V – total volume of the culture broth, [l];
$D_L^{IN} = F_{IN}/V$ – dilution level, [h^{-1}];
$D_L^{OUT} = F_{OUT}/V$ – dilution level of dropping, [h^{-1}];
F_{IN} – feeding rate, [l/h];
F_{OUT} – dropping rate, [l/h].

The Levenberg-Marquardt algorithm with least squares objective function is used for parametric optimization.

Table 4. Final estimation of the model parameters

Estimates		
K_1 = 0,73081783161393	K_2 = 0,813890332396588	K_3= 1,176543
a_0 = - 9,94668760424156	c_0 = -24,3981562776857	d_0 = - 10,6162340651542
a_1 = 0,00890419741892455	c_1 = 49,4521101341518	d_1 = 46,1913193447749
a_2 = - 1,35913448319213.10^{-6}	c_2 = - 41,9735	d_2 = - 127,521
b_1 = 34,0537484058599	c_3 = 30,9143	d_3 = 123,7073
b_2 = - 26,0546580593866	c_4 = 2,00615	d_4 = 0,36768
b_3 = - 0,0364334337803236	c_5 = - 0,0648347	d_5 = - 0,0017355
	c_6 = 0,000724185	d_6 = - 0,000075946
	c_7 = - 0,735849	d_7 = - 0,407434

During the parametric optimization experimentally established optimal modes of the feeding rate, dropping rate and oxygen saturation rate are applied. The simulation results are shown in Figs. 6, 7, 8, 9.

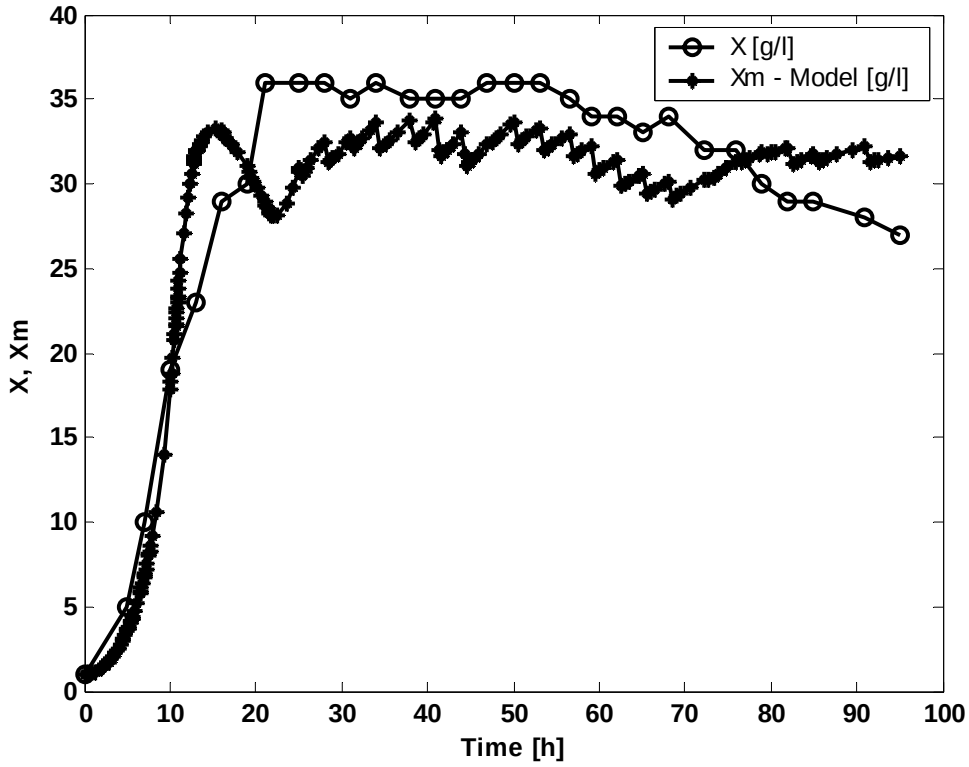

Fig. 6 Time course of the biomass concentration

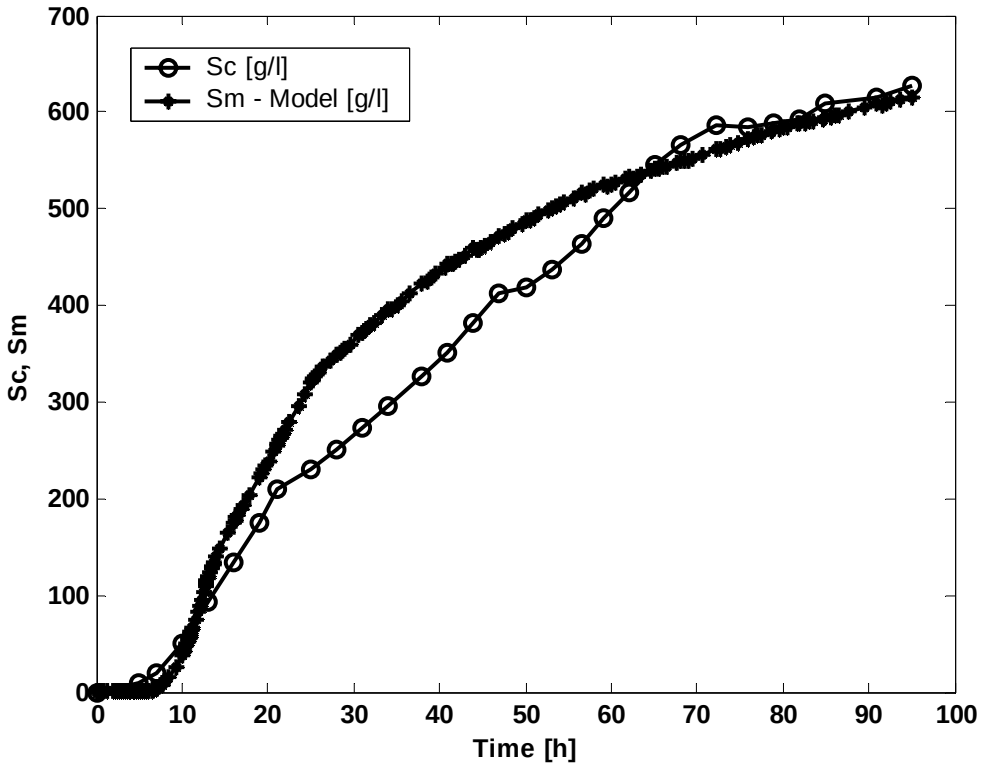

Fig. 7 Time course of the consumed substrate

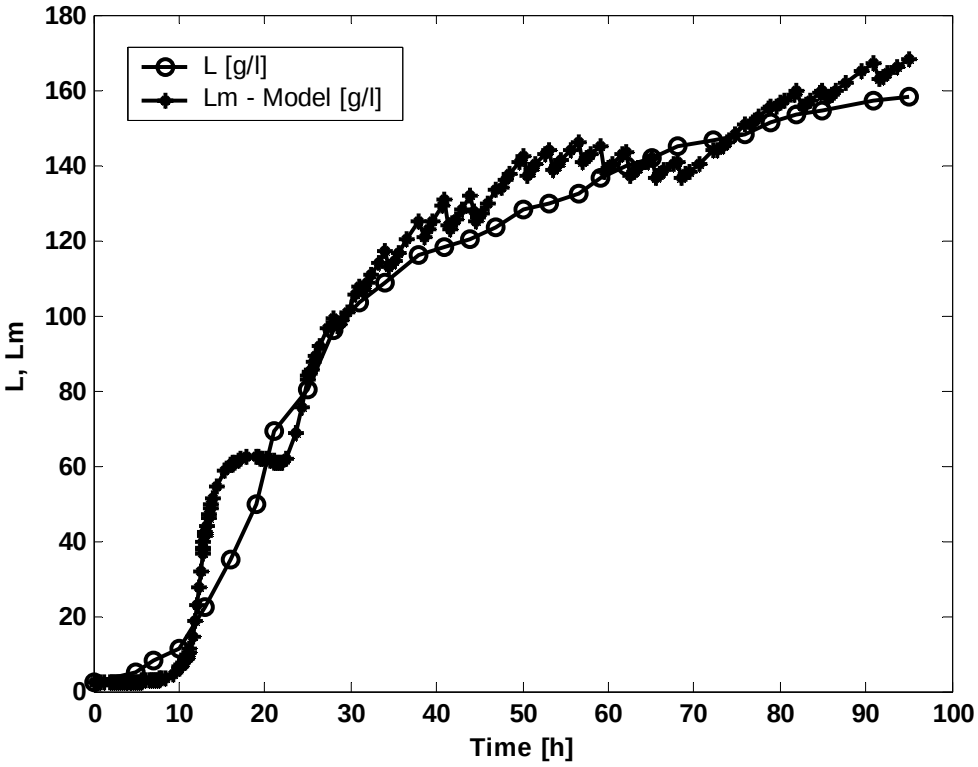

Fig. 8 Time course of the L-lysine production

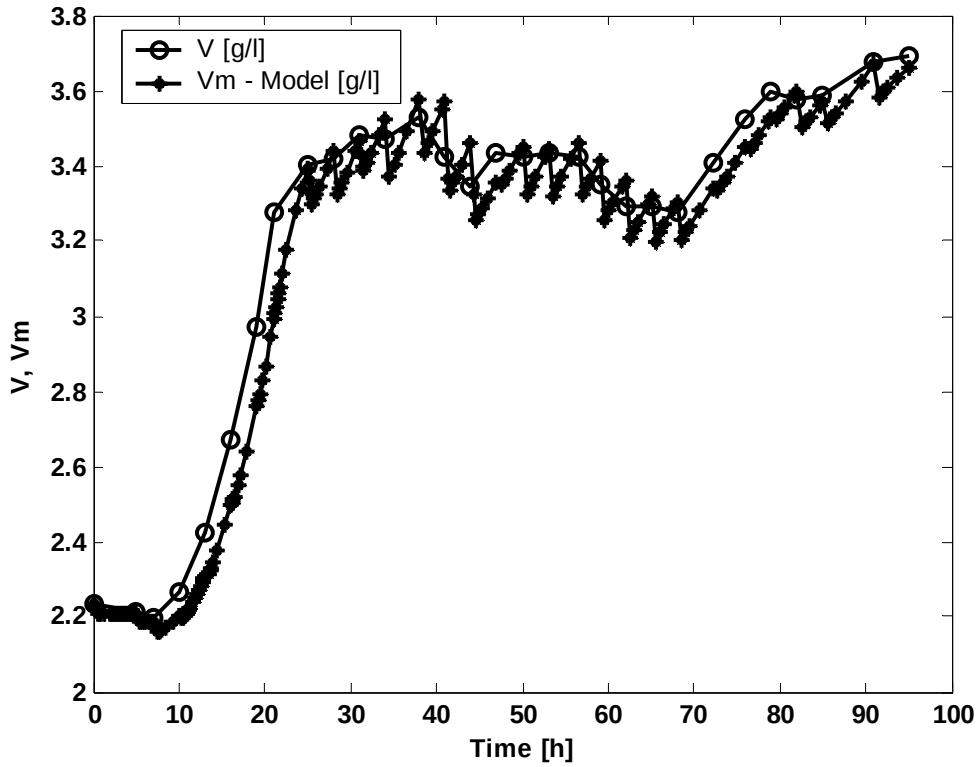

Fig. 9 Time course of the total volume of the culture broth

The presented simulation results confirmed the adequacy oh the presented model.

Conclusions

The following conclusions could be drown based on the results achieved so far:

1. The trend and values of the specific rates are estimated based on the experimental data and material balance followed by an additional data processing.
2. The linear regression or polynomial regressions are applied for selection of a preliminary structure of the models describing the specific rates and initial estimates of their parameters. The aim is the selection of the appropriate model structure and the model fit to the experimental data. The full regression analysis is done including the investigation of the residuals.
3. Non-linear regression, based on the selected model structure and initial values of the parameters, without any data transformations is applied as a next stage for mathematical model development. The model selection is done using R^2 coefficient and the results of the residual investigation.
4. The final stage of the investigation is connected with the parametric optimization of the model through the non-linear optimization procedure under the confidence intervals of the parameters using Optimization Toolbox. The Levenberg-Maquardt algorithm with least squares objective function is used for optimization.
5. Based on the simulation results it could be concluded that the obtained mathematical model describes the trend of the experimental data in a satisfactory way.

References

1. Bastin G., L. Chen, V. Chotteau (1992). Can we Identify Biotechnological Processes?, Proceedings of IFAC, Modelling and Control of Biotechnological Processes, Colorado, USA, 83-88.
2. Georgiev Tz., Al. Ratkov, St. Tzonkov (1997). Mathematical Modelling of Fed-batch Fermentation Processes for Amino Acid Production, Mathematics and Computers in Simulation, 44, 171-285.
3. Georgiev Tz., Al. Ratkov, St. Tzonkov (1995). An Approach for Mathematical Modelling of Fed-batch Process for L-lysine Production, Biotechnology and Biotechnological equipment, 4, 84-92.
4. Leuchtenberger W., K. Huthmacher, K. Drauz (2005). Biotechnological Production of Amino Acids and Derivatives: Current Status and Prospects, Appl. Microbiol. Biotechnol., 69, 1-8.
5. MathWorks Inc. (2003). MATLAB User's Guide.
6. MathWorks Inc. (2003). System Identification Toolbox, User's Guide.
7. MathWorks Inc. (2003). SIMULINK, Using SIMULINK.
8. MathWorks Inc. (2003). Optimization Toolbox, User’s Guide.
9. STATGRAPHICS (2000). Version 4.0 Plus For Windows, User Manual, Magnugistics Inc. USA.