Bioautomation, 2007, 6, 17 – 26 ISSN 1312 – 451X

One Approach for Dynamic Modelling of L-lysine Repeated Fed-batch Fermentation

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

[1] *Institute of Microbiology, Bulgarian Academy of Science*
26 Acad. G. Bonchev Str., Sofia 1113, 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
Sofia 1756, Bulgaria, Phone: (+359 2) 965 2408
E-mail: tzg@tu-sofia.bg

[*] *Corresponding author*

Received: December 20, 2006 **Accepted: :Mart 2, 2007**

Published: Mart 27, 2007

Abstract: *This article deals with establishment of dynamic unstructured model of variable volume fed-batch fermentation process with intensive droppings for L-lysine production. The presented approach of the investigation includes the following main procedures: description of the process by generalized stoichiometric equations; preliminary data processing and calculation of specific rates for main kinetic variables; identification of the specific rates as a second-order non-linear dynamic models; establishment and optimisation of dynamic model of the process; simulation researches. MATLAB is used as a research environment.*

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

Introduction

In terms of market volume, development over the last 20 years has been tremendously bullish in the so-called feed amino acids L-lysine, DL-methionine, L-threonine, and L-tryptophan. These four amino acids constitute the largest share (56%) of total amino acid market, estimated in 2004 at approximately US $ 4.5 billion [5]. Bioprocess technologies for their production were significantly improved by application of more sophisticated feeding and automation strategies [4].

The synthesis of mathematical models for biotechnological processes in principal is known to be the major task for the application of modern control science for their optimization. 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].

Some approaches for dynamic modelling of L-lysine fed-batch fermentation process have been described in our previous articles [2, 3] including modelling of fed-batch fermentation process with droppings [10]. This article aims to present another approach for development of dynamic unstructured model for L-lysine fed-batch fermentation process with intensive droppings of the culture broth as well as investigation of the specificity of the process and its reflection on the obtained mathematical model. This approach presents extended variant of

generalized stoichiometric equations takes in to account the dropping conditions by sequential reactions. The specific rates are described by additional sequential reactions with appropriate order. Second-order non-linear dynamic models of the specific rates: growth rate (μ), substrate utilization rate (ν), production rate (ρ) are included in this approach of investigation based on preliminary data processing and two stage identification procedure.

Identification procedure, applied for estimation of the model structure and coefficients, taking in consideration the specificity concerning dropping procedure. The important stage of this procedure is the parametric optimisation of the model. The procedures for identification, optimisation and simulation researches are realized by **MATLAB** and **STAGRAPHICS** packages [6, 7, 8, 9, 11]. Main approaches and steps, used for development of mathematical models, are described in details in our previous articles [2, 3, 10].

Materials and methods

Experimental procedures and analysis

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 with prototrophic nature, which is insuring successful carrying of fed-batch process with big number of droppings. Analytical methods used are as follows: biomass is measured as dry cell mass, [g/l]; sugar concentration – as reducible compounds, [g/l]; L-lysine – by chromatographic method, [g/l]. During the process *on-line* measurement of differed physical-chemical variables have been measured by proper sensors.

The experimental data are shown in Fig. 1. Dissolved oxygen tension [%] is denoted as DO in Fig. 1.

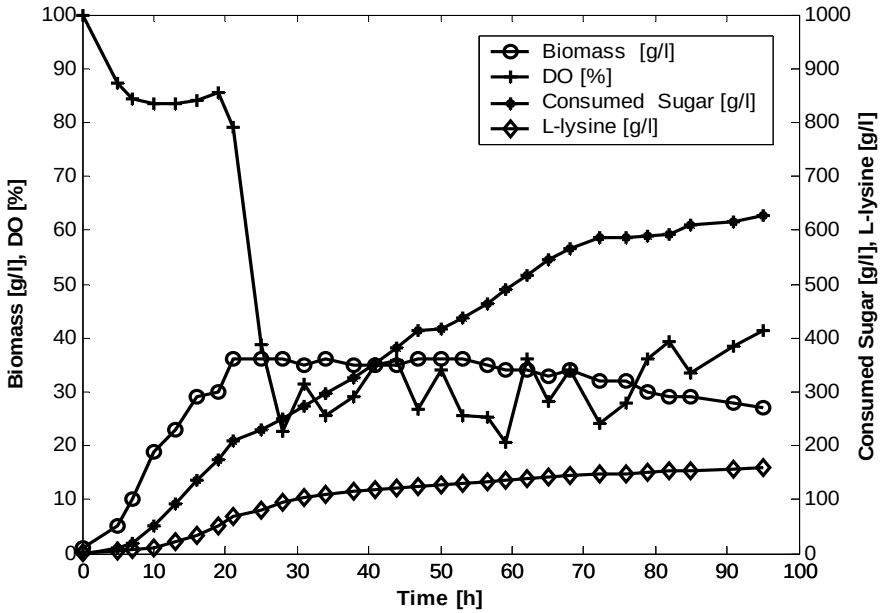

Fig. 1 Time course of the experimental data

Primary data processing

The primary data processing of the experimental data includes the following stages: transformation the different measurable units to unit [d/l] taking into account changes (as

strict balance) of the working volume during the process; calculation of the specific rates: growth rate (μ) [h^{-1}], substrate utilization rate (ν) [h^{-1}], production rate (ρ) [h^{-1}]. To reach the final aim of primary data processing - calculation of the specific rates – the appropriate optimal approximation of the experimental data is done by a spline functions. The derivatives of variables describing the process kinetics are also calculated. The specific growth rate, the rate of substrate utilisation, the specific production rates are calculated by the equations:

$$\hat{\mu} = \frac{\dot{X}_t}{X_t}, \hat{\nu} = \frac{\dot{S}_c}{X_t}, \hat{\rho} = \frac{\dot{L}_t}{X_t}, \quad (1)$$

where: X_t – total biomass produced at proper time expressed as [g/l], S_c – total sugar consumed at proper time expressed as [g/l], L_t – total L-lysine produced at proper time expressed as [g/l] and derivatives of these variables denotes as follows $\dot{X}_t$, $\dot{S}_c$, $\dot{L}_t$. The specific rates are shown in Fig. 2.

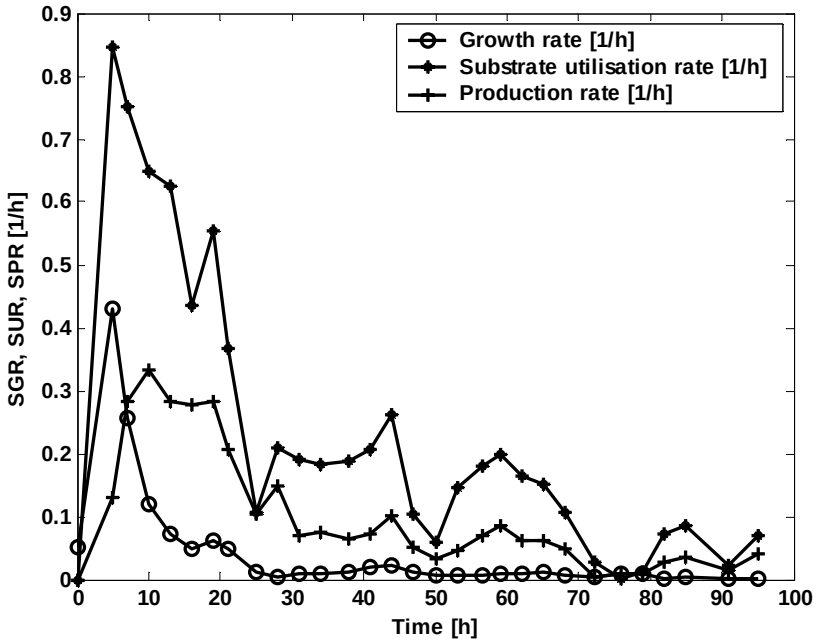

Fig. 2 Time course of the specific rates
The notations are: SGR – specific growth rate, SUR – specific sugar utilisation rate, SPR – specific lysine production rate.

Results and discussion

The identification procedure includes the following main stages: determination of the set of generalized stoichiometric equations; identification of the specific rates; description of the dynamic model subject to dropping conditions; parametric optimisation of the obtained model and simulation.

Generalized stoichiometric equations

The fermentation processes could be described by the following scheme of generalized stoichiometric equations:

$$\begin{aligned}
&\mathbf{S_C} \xrightarrow{\varphi_X} \mathbf{X} \\
&\mathbf{C} + \mathbf{S_C} \xrightarrow{\varphi_g} \mathbf{X} \xrightarrow{\varphi_{OUT}} \mathbf{X} \\
&\mathbf{S_R} \xrightarrow{\varphi_S} \mathbf{S_C} \xrightarrow{\varphi_{OUT}} \mathbf{S_C} \quad , \\
&\mathbf{C} + \mathbf{S_C} + \mathbf{X} \xrightarrow{\varphi_L} \mathbf{L} \xrightarrow{\varphi_{OUT}} \mathbf{L} \\
&\mathbf{V_0} \xrightarrow{\varphi_F} \mathbf{V_f} \xrightarrow{\varphi_{OUT}} \mathbf{V}
\end{aligned} \tag{2}$$

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

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

$$\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}(\mathbf{S_C}, \mathbf{C}), \hat{\mathbf{v}} = \hat{\mathbf{v}}(\hat{\boldsymbol{\mu}}, \mathbf{X}), \hat{\boldsymbol{\rho}} = \hat{\boldsymbol{\rho}}(\hat{\boldsymbol{\mu}}, \mathbf{X}) \tag{3}$$

The proposed hypothesises are discussed below. It could be seen (Fig. 2) that some of the fundamental kinetic relations, for example Monod law, are not performed for the obtained specific rates. On the other hand the fundamental relations such a Monod law reveals the approach to reach new relations. The Monod law has the following properties: hidden linear structure (the reciprocal transformation reveals this structure); hidden exponential structure if the Monod law is divided by the substrate concentration (**S**) the result reveals Pade' approximant of the function $\mathbf{EXP(-K_S/S)}$. The second hidden property of the Monod law gives an argument for the exponential structure investigation of the specific rates under hypotheses (3).

The rates of the generalized stoichiometric reactions are revealed as follows:

$$\begin{aligned}
&\varphi_X = \varphi_G = \hat{\boldsymbol{\mu}}(\mathbf{S_C}, \mathbf{C})\mathbf{X}, \\
&\varphi_S = \hat{\mathbf{v}}(\hat{\boldsymbol{\mu}}, \mathbf{X})\mathbf{X}, \ \varphi_L = \hat{\boldsymbol{\rho}}(\hat{\boldsymbol{\mu}}, \mathbf{X})\mathbf{X}, \\
&\varphi_F = \mathbf{D_L^{IN}V}, \ \varphi_{OUT} = \mathbf{D_L^{OUT}V},
\end{aligned} \tag{4}$$

where: $\mathbf{D_L^{IN} = F_{IN}/V}$ – dilution level, [h^{-1}]; $\mathbf{D_L^{OUT} = F_{OUT}/V}$ – dilution level as a result of droppings, [h^{-1}]; $\mathbf{F_{IN}}$ – feeding rate, [h^{-1}]; $\mathbf{F_{OUT}}$ – dropping rate, [h^{-1}].

The process with droppings is very complex one. There are a couple of phases as: batch phase, fed-batch phase and fed-batch with dropping. Based on our experience we suppose that the sequence of phases could be described by the dynamic models of the all specific rates. Generalized sequential reactions (4) help as to construct the dynamic models of the specific rates.

Suppose that the specific rates could be described by generalized stoichiometric equations as follows:

$$\begin{aligned}
&\mathbf{X} \xrightarrow{\varphi_{\mu 1}} \boldsymbol{\mu}_1 \xrightarrow{\varphi_{\mu 2}} \boldsymbol{\mu}_2 \xrightarrow{\mathbf{K}_\mu} \mathbf{X} \\
&\mathbf{S}_\mathbf{C} \xrightarrow{\varphi_{\nu 1}} \boldsymbol{\nu}_1 \xrightarrow{\varphi_{\nu 2}} \boldsymbol{\nu}_2 \xrightarrow{\mathbf{K}_\nu} \mathbf{S}_\mathbf{C}, \\
&\mathbf{L} \xrightarrow{\varphi_{\rho 1}} \boldsymbol{\rho}_1 \xrightarrow{\varphi_{\rho 2}} \boldsymbol{\rho}_2 \xrightarrow{\mathbf{K}_\rho} \mathbf{L}
\end{aligned} \qquad (5)$$

where: $\varphi_{\mu 1}$, $\varphi_{\mu 2}$, $\varphi_{\nu 1}$, $\varphi_{\nu 2}$, $\varphi_{\rho 1}$, $\varphi_{\rho 2}$, $\mathbf{K}_\mu$, $\mathbf{K}_\nu$, $\mathbf{K}_\rho$ are rates of the generalized stoichiometric reactions for the specific tares.

The system of the stoichiometric reactions (5) describes a hypothesis that the specific rates include "hidden" states. These states are denoted by two variables (μ_1, μ_2, ν_1, ν_2, ρ_1, ρ_2). The first one is used to describe the specific rate (growth, substrate utilization or production). The second one is used to describe the derivative of the specific rate with respect to time (derivatives of the growth, substrate utilization or production rates).

Dynamic model of the fed-batch process with dropping

The dynamic model subject to material balance of the process could be described as follows:

$$\frac{\mathbf{d}\boldsymbol{\xi}}{\mathbf{dt}} = \mathbf{K}\boldsymbol{\varphi}(\boldsymbol{\xi}) - \mathbf{D}_\mathbf{L}^{\mathbf{IN}}\boldsymbol{\xi} + \mathbf{F} - \mathbf{D}_\mathbf{L}^{\mathbf{OUT}}\boldsymbol{\xi}, \qquad (6)$$

where $\boldsymbol{\xi}^\mathbf{T} = [\mathbf{X}, \mathbf{Sc}, \mathbf{L}, \mathbf{V}]^\mathbf{T}$ – state space vector; $\boldsymbol{\varphi}^\mathbf{T} = [\tilde{\boldsymbol{\mu}}\mathbf{X}, \tilde{\boldsymbol{\nu}}\mathbf{X}, \tilde{\boldsymbol{\rho}}\mathbf{X}, \varphi_\mathbf{F}]^\mathbf{T}$ – vector of reaction's rates; $\mathbf{F}^\mathbf{T} = [\mathbf{0}, \mathbf{f}_1, \mathbf{0}, \mathbf{0}]^\mathbf{T}$ – vector of input streams with elements $\mathbf{f}_1 = \mathbf{D}_\mathbf{L}^{\mathbf{IN}}\mathbf{S}_\mathbf{IN}$, matrix $\mathbf{K}$ with yield coefficients $\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_3$.

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{2} \end{bmatrix}$$

Suppose the specific rates could be described by second-order differential equations subject to stoichiometric equations (5) as:

$$\begin{aligned}
&\frac{\mathbf{d}\boldsymbol{\mu}}{\mathbf{dt}} = \mathbf{M}\boldsymbol{\mu} + \mathbf{m}\hat{\boldsymbol{\mu}}(\mathbf{S}_\mathbf{C}, \mathbf{C}) \\
&\tilde{\boldsymbol{\mu}} = \mathbf{L}\boldsymbol{\mu},
\end{aligned} \qquad (7a)$$

where: $\boldsymbol{\mu}^\mathbf{T} = [\mu_1, \mu_2]^\mathbf{T}$ – state space vector of the specific growth rate, $\mathbf{m}^\mathbf{T} = [\mathbf{0}, \mathbf{m}_0]^\mathbf{T}$, $\mathbf{L} = [\mathbf{1}, \mathbf{0}]$, $\mathbf{M}$ – matrix as follows:

$$\mathbf{M} = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{m}_1 & -\mathbf{m}_2 \end{bmatrix},$$

$$\begin{aligned}
&\frac{\mathbf{d}\boldsymbol{\nu}}{\mathbf{dt}} = \mathbf{N}\boldsymbol{\nu} + \mathbf{n}\hat{\boldsymbol{\nu}}(\boldsymbol{\mu}, \mathbf{X}), \\
&\tilde{\boldsymbol{\nu}} = \mathbf{L}\boldsymbol{\nu},
\end{aligned} \qquad (7b)$$

where: $\mathbf{v}^T = [\mathbf{v}_1, \mathbf{v}_2]^T$ – state space vector of the specific growth rate, $\mathbf{n}^T = [0, n_0]^T$, $\mathbf{L} = [1, 0]$, $\mathbf{N}$ – matrix as follows:

$$\mathbf{N} = \begin{bmatrix} 0 & 1 \\ -n_1 & -n_2 \end{bmatrix},$$

$$\frac{d\boldsymbol{\rho}}{dt} = \mathbf{R}\boldsymbol{\rho} + \mathbf{r}\hat{\rho}(\boldsymbol{\mu}, \mathbf{X}), \qquad (7c)$$
$$\tilde{\rho} = \mathbf{L}\boldsymbol{\rho},$$

where: $\boldsymbol{\rho}^T = [\rho_1, \rho_2]^T$ – state space vector of the specific growth rate, $\mathbf{r}^T = [0, r_0]^T$, $\mathbf{L} = [1, 0]$, $\mathbf{R}$ – matrix as follows:

$$\mathbf{R} = \begin{bmatrix} 0 & 1 \\ -r_1 & -r_2 \end{bmatrix},$$

where: m_0, m_1, m_2, n_0, n_1, n_2, r_0, r_1, r_2 are coefficients.

Estimation of the initial values of the coefficients and initial conditions of the system (7a) - (7c) are very important for modelling of the specific rates.

Dropping conditions

It is assumed that at the discrete time moments of the dropping the derivatives of the kinetic variables are equal to zero. The dropping conditions subject to model (6) are satisfied as follows:

- Dropping conditions for growth

$$\mathbf{F}_{OUT} = \mathbf{K}_1 \boldsymbol{\mu} \mathbf{V}(t_k) - \mathbf{F}_{IN} \qquad (8)$$

- Dropping conditions for L-lysine production

$$\mathbf{F}_{OUT} = \mathbf{K}_3 \boldsymbol{\rho} \frac{\mathbf{X}(t_k)}{\mathbf{L}(t_k)} \mathbf{V}(t_k) - \mathbf{F}_{IN} \qquad (9)$$

- Dropping conditions for substrate utilization

$$\mathbf{F}_{OUT} = \mathbf{K}_2 \mathbf{v} \frac{\mathbf{X}(t_k)}{\mathbf{S}(t_k)} \mathbf{V}(t_k) + \mathbf{F}_{IN} \left(\frac{\mathbf{S}_{IN} - \mathbf{S}_C(t_k)}{\mathbf{S}_C(t_k)} \right) \qquad (10)$$

The comparison of the expressions (9) - (10) obtains the equality

$$\mathbf{K}_3 \boldsymbol{\rho} \frac{\mathbf{X}(t_k)}{\mathbf{L}(t_k)} = \mathbf{K}_2 \mathbf{v} \left(\frac{\mathbf{X}(t_k)}{\mathbf{S}_C(t_k)} \right) + \frac{\mathbf{F}_{IN}}{\mathbf{V}} \left(\frac{\mathbf{S}_{IN}}{\mathbf{S}_C(t_k)} \right) \qquad (11)$$

It could be emphasized that these conditions are satisfied at the discrete time moments. The specific rates included in the above equalities are the solution of the system (7a) - (7c).

Models of the specific rates

The identification procedure of the specific rates includes the following steps:

- Application of linear regression for selection of suitable set of predictors and initial estimates of the parameters
- Non-linear regression application for final parameter estimation.

The results of the identification procedure are derived by **STATGRAPHICS** facilities as follows.

Table 1. Models of the specific rates

Model	R^2
$\hat{\mu} = EXP(a_0 + a_1 S_C + a_2 S_C^2 + b_1 C + b_2 C^2 + b_3 (S_c C))$	$R^2 = 0.89$ (12)
$\hat{v} = EXP(c_0 + c_1 \mu + c_2 \mu^2 + c_3 \mu^3 + c_4 X + c_5 X^2 + c_6 X^3 + c_7 (\mu X))$	$R^2 = 0.81$ (13)
$\hat{\rho} = EXP(d_0 + d_1 \mu + d_2 \mu^2 + d_3 \mu^3 + d_4 X + d_5 X^2 + d_6 X^3 + d_7 (\mu X))$	$R^2 = 0.84$ (14)

The acceptation of the specific rate models is based on the value of the determination coefficient (R^2) obtained by non-linear regression. Investigation of the residuals shows that the majority parts of them are normally distributed but the serial correlation could be seen.

Optimization and simulation

The second stage of this procedure is connected with the parametric optimization of the model. The results are shown in Table 2.

Table 2. Estimated parameters

$K_1 = 1.0864$	$c_6 = -0.0025944$
$a_0 = -10.8514$	$c_7 = 0.9265999$
$a_1 = 0.009682$	$n_0 = 0.0604141$
$a_2 = -1.7746762*10^{-6}$	$n_1 = 0.0164165$
$b_1 = 34.131772$	$n_2 = 0.1209947$
$b_2 = -25.90943$	$K_3 = 0.6308375$
$b_3 = -0.03753$	$d_0 = -9.593265$
$m_0 = 1.0435$	$d_1 = 41.8108988$
$m_1 = 1.239878$	$d_2 = -38.3577$
$m_2 = 0.586369$	$d_3 = -41.07372$
$K_2 = 0.650056$	$d_4 = 0.7945639$
$c_0 = -24.2277$	$d_5 = -0.0365328$
$c_1 = 56.37046$	$d_6 = 0.00054084$
$c_2 = -516.2815$	$d_7 = -0.7418933$
$c_3 = -290.2391$	$r_0 = 1.8445898$
$c_4 = -1.770672$	$r_1 = 1.550243$
$c_5 = 0.157532$	$r_2 = 8.658613$

During the parametric optimization the experimentally established optimal modes of the feeding rate and oxygen saturation are applied. The concentration S_{IN} as a function is also obtained by an optimization based on the experimental data. The results are shown below.

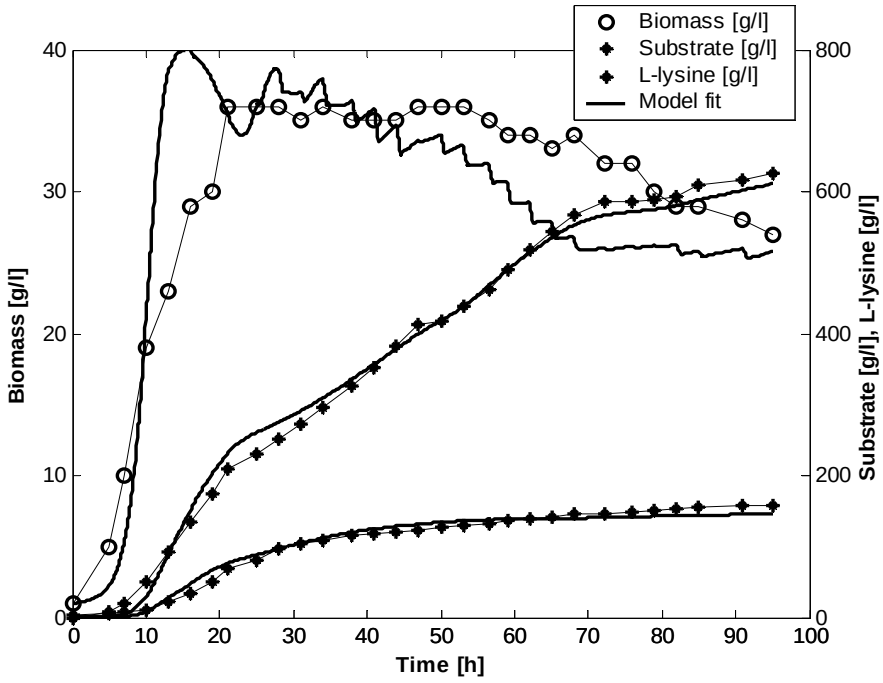

Fig. 3 Time course of the experimental data and model for main kinetic variables

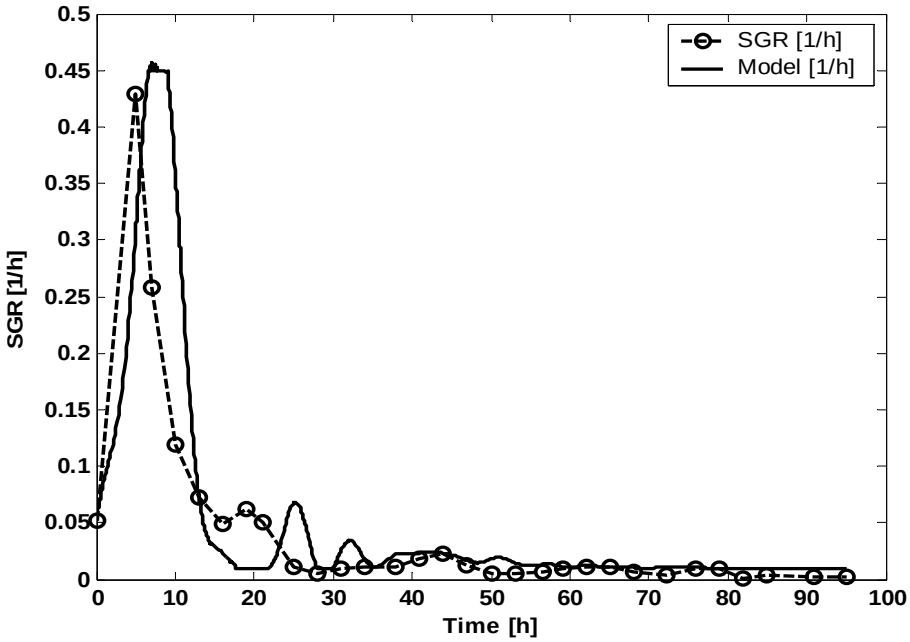

Fig. 4 Model approximation fit of the specific growth rate
(SGR – specific growth rate)

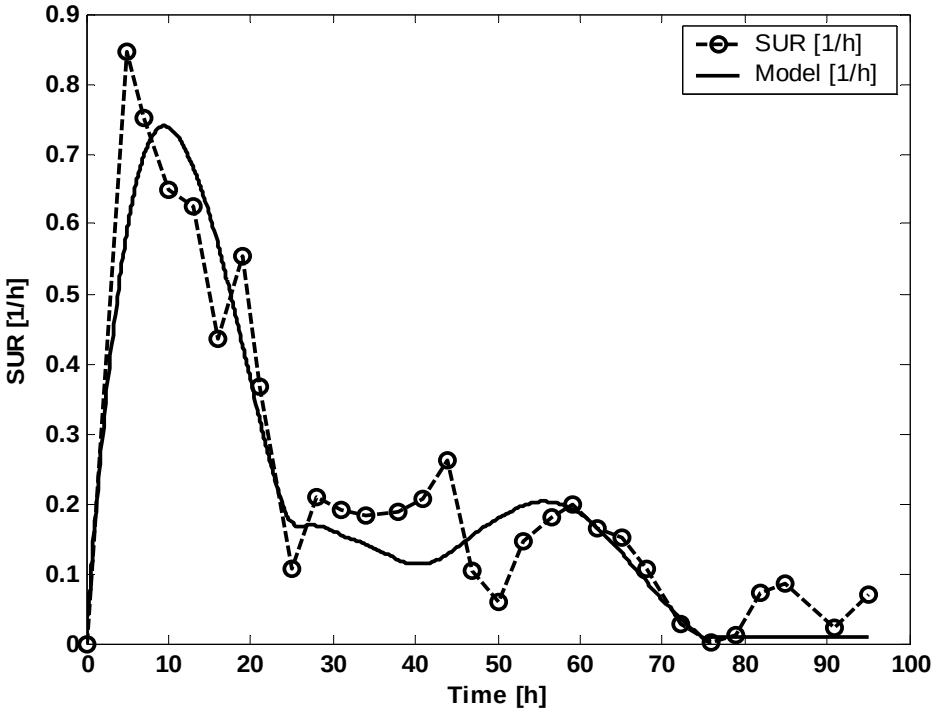

Fig. 5 Model approximation fit of the specific sugar utilisation rate (SUR – specific sugar utilisation rate)

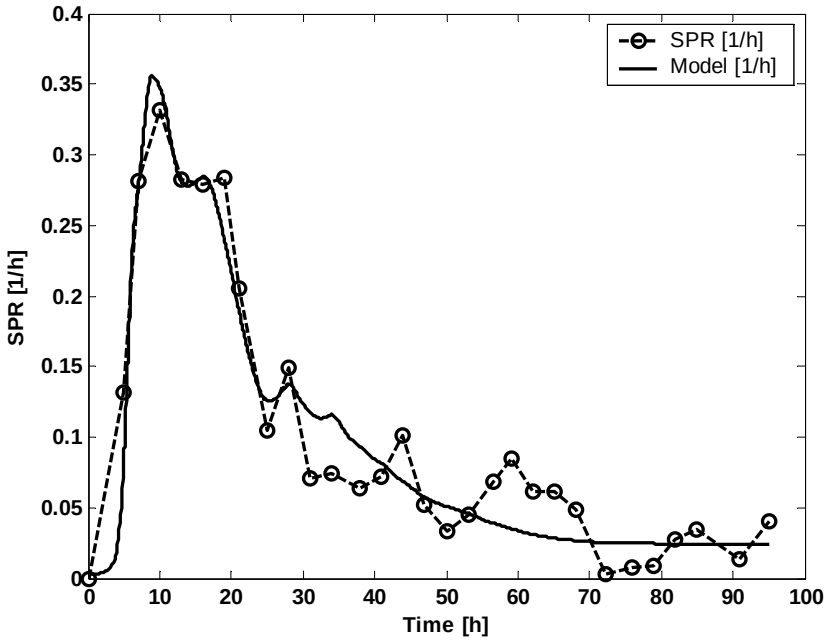

Fig. 6 Model approximation fit of the specific lysine production rate (SPR – specific lysine production rate)

Conclusions

The following consideration and conclusions can be derived from the investigation, connected with development of mathematical model of the discussed process.

1. The trend and values of the specific rates are estimated based on the experimental data and material balance followed by 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 its parameters. The aim is 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 is applied without any data transformations as a next stage for mathematical model development. Using $\mathbf{R^2}$ coefficient and the results of the residual investigation does the model selection.
4. For more precise description dynamic non-linear models of the specific rates have been developed.
5. The final stage of the investigation is connected with the parametric optimisation of the model through the non-linear optimisation procedure under the confidence intervals of the parameters using Optimisation Toolbox. The Levenberg-Maquardt algorithm with least squares objective function is used for optimisation.
6. The mathematical model describes the trend of the newly investigated experimental data and specific rates of the main kinetic variables.

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. Herman T. (2003). Industrial Production of Amino Acids by Coryneform Bacteria, Journal of Biotechnology, 104, 155-172.
5. Leuchtenberger W, K. Huthmacher, K. Drauz (2005). Biotechnological Production of Amino Acids and Derivatives: Current Status and Prospects, Appl. Microbiol. Biotechnology, 69, 1-8.
6. MathWorks Inc. (1996). MATLAB User's Guide.
7. MathWorks Inc. (1996). System Identification Toolbox, User's Guide.
8. MathWorks Inc. (1998). SIMULINK, Using SIMULINK.
9. MathWorks Inc. (1996). Optimisation Toolbox, User's Guide.
10. Ratkov Al., Georgiev Tz., Ivanova V., Kristeva J. (2004). Investigation of L-lysine Fed-batch Fermentation Process with Droppings by Mathematical Modeling, – Proceedings of ITI'2004, 26th International Conference of Information Technology Interfaces, June 7-10, Cavtat/Dubrovnic, Croatia, 211-216
11. STATGRAPHICS (1995). Version 4.0 Plus For Windows, User Manual, Magnugistics Inc. USA.