MULTIMODEL IDENTIFICATION OF BIOTECHNOLOGICAL PROCESS

Yuliana Hristova Hristova[1], **Svetla Ivanova Vassileva**[2]

[1]Technical college – Yambol, Graf Ignatiev Str.,38,8600 Yambol, e-mail:yhristova@tk.uni-sz.bg
[2]Institute of Control and Systems Research - BAS, Acad.G.Bonchev Str., bl.2, 1113 Sofia, e-mail: vasileva@icsr.bas.bg

Dynamical behavior of complex nonlinear processes could be presented as a sequence of phases. The methodology, which can provide their adequate mathematical description is based of synthesis of local models, for which is designed appropriate control low.

The basic purposes of the presented paper are identification and synthesis of multimodel control for nonlinear biotechnological process.

Here is done a modeling and validation parameters of kinetic model of biotechnological process with the help of Kalman filter (fig.1). In the paper is calculated an assessment of the specific growth rate. The next stage of the work aims defining of separate phases of the studying process, application of Kalman filtration for each of them and on this base accomplishment of comparison of the accuracy at work without dividing of process of phases with multi phase process presentment.

Keywords: multimodel identification, Kalman filter, biotechnological process

1. INTRODUCTION

The specific growth rate is one of the most important parameters in biotechnological processes. The relationships between the specific rates of growth, substrate consumption, and product formation are crucial for monitoring and controlling these processes. Common practice in specific growth rate determination is fitting exponential or polynomial functions through the biomass measurement results. Apart from the unknown order of the polynomial function, the main disadvantage of such an approach is the assumption that the specific growth rate remains constant or follows a polynomial pattern. In batch and fed-batch processes, substrate and biomass concentrations change and as a consequence the specific growth rate is subject to changes. Therefore, it is of importance to be able to monitor the specific growth rate as a function of time without a priori knowledge of the order in which it changes.[1]

There are great number of reports about recursive estimation of biological processes [4,6,7]. Several different techniques are used amongst such as the extended Kalman filter and the asymptotic observer. These are intended as software sensors for online use. On them each estimation is based only on the measurements up to the time under consideration and as a result the estimators react too late to changes.

There are two main approaches for on-line state and parameter estimation of nonlinear systems such as biotechnological processes – exponential estimators design (based on Kalman filtering method) and asymptotic estimators design (based on a linear algebra results) [4].

Kalman filtering method leads to exponential observers for which it is possible to arbitrarily fix the speed of convergence of the estimated variables towards their true values. Much of the work concentrated on Kalman filtering approach dealt with a model, the accuracy of which played a major determining factor in the quality of estimation. Deficiencies in the process kinetic model, such as unmodelled process dynamics and model parameters' variations, have to a certain extent been overcome by applying adaptive estimation techniques, where the process state and some model parameters are simultaneously estimated [2].

This paper describes a modelling technique of biotechnological process based on the extended Kalman filter - single phased and multiphased.

2. MATERIALS AND METHODS

Methods for growth rate estimation

Traditional methods

For model development one of the main tasks is to reconstruct the time function of the specific growth rate, ● (t) from biomass measurements, X(t). Besides dividing the total growth rate by the average biomass concentration between two measurements, biotechnology literature [1], present the following frequently used function to calculate the specific growth rate between two consecutive biomass measurements:

$$\mu = \frac{\mathbf{ln}\left(\frac{X_{k+1}}{X_k}\right)}{t_{k+1} - t_k} \qquad (1)$$

With this approach, the specific growth rate is calculated over each time interval; however, with a decreasing time interval, noise on the biomass measurements plays an increasing role. Other approaches seen regularly are fitting exponential curves or other [1]. Main disadvantage of fitting such equations through data is the assumption that the specific growth rate is constant and so these methods are not appropriate to determine changes. Polynomials or splines solve that problem, but fix the specific growth rate to change according to a polynomial of a lower degree. The main disadvantage of all these traditional techniques is that they consider each measurement to be exact by disregarding the accuracy and noise characteristics of the biomass measurements.

Recursive estimation

Several authors described recursive estimators for determination of specific growth rate [1] and [6]. The recursive estimator uses a model to reconstruct the state when only part of the state is measured. The model necessary for specific growth rate estimation describes changes of biomass as directly related to the biomass concentration, net specific growth rate and the dilution rate, D (for a batch biotechnological process D=0):

$$\frac{dX}{dt} = \mu X - DX \qquad (2)$$

The behavior of the specific growth rate during cultivation is not known a priori, that why the change of the specific growth rate is assumed to be a random walk process. Since all unpredicted changes of the biomass have to be reflected by the specific growth rate, the model description for biomass growth is assumed perfect.

Extended Kalman Filter (EKF)

The EKF algorithm, a frequently used recursive estimator [1], is a two-step algorithm using a time-update and a measurement-update. Given the current estimate, $\hat{x}_k$, the time update predicts the state value at the next sample, $\bar{x}_{k+1}$. Here is executed a one-step-ahead prediction using Runge-Kutta integration. The prediction of the corresponding error covariance, $\bar{P}_{k+1.}$, is calculated using the linearised model. The next step in the EKF algorithm, the measurements-update, consists of approximation of the error covariance, $\hat{P}_{k+1}$ and the correction of the state based on the new measurement, y_{k+1}. There is used a correction term, which is a function of the innovation, that is, the discrepancy between the measured and predicted values.

In contrast to the traditional techniques, the Kalman filter algorithm takes the measurement error characteristics into account. The EKF can be implemented online since it is based on the measurements up to and including the moment of estimation.

3. EXAMPLE

In the present work Matlab and Simulink (Mathworks, Mass., USA) were used for the development of the extended Kalman filter algorithms, the simulation of microbial growth, and for the visualisation of the results.

The following dynamic model describes batch process of glucoamylase production:

$$\frac{dX}{dt} = \mu X \qquad (3)$$

$$\frac{dS}{dt} = -Y_{S/X} X \qquad (4)$$

$$\frac{dP}{dt} = \eta X, \qquad (5)$$

where S is the substrate concentration (maltodextrine and corn steep liguor) [g/l], P is the glucoamylase concentration [g/l], X is the mycelium biomass concentration [g/l], $Y_{S/X}$ is the yield of biomass depending on substrate utilization, η is the yield of glucoamylase concerning the biomass formation.

The specific growth rate of mycelium biomass is:

$$\mu = \frac{\mu_{\mathbf{max}} S}{K_S + S}, \qquad (6)$$

where $\mu_{\mathbf{max}}$ is the maximal specific growth rate [1/h], K_s is the Michaelis Menten constant [g/l].

In this work, the Kalman observer is applied for estimation of the model parameters on the base of measurement the state variables - $X(t)$, $S(t)$ $P(t)$. Its design is based on a "quadratic optimization" approach [5].

A set of experimental data for biomass, X, substrate concentration, S, and product P is applied in order to investigate the system dynamics in the presence of the designed estimator.

As a next stage of the problem of estimation model parameters in the present work are defined a separate phases of the studying nonlinear process and is applied a Kalman filter for each of them. The purpose is to compare an accuracy of single-phase modeling and multi model process presentation.

The results are presented in Fig.1-Fig.4. The estimated values are shown by line and experimental data by dotted line.

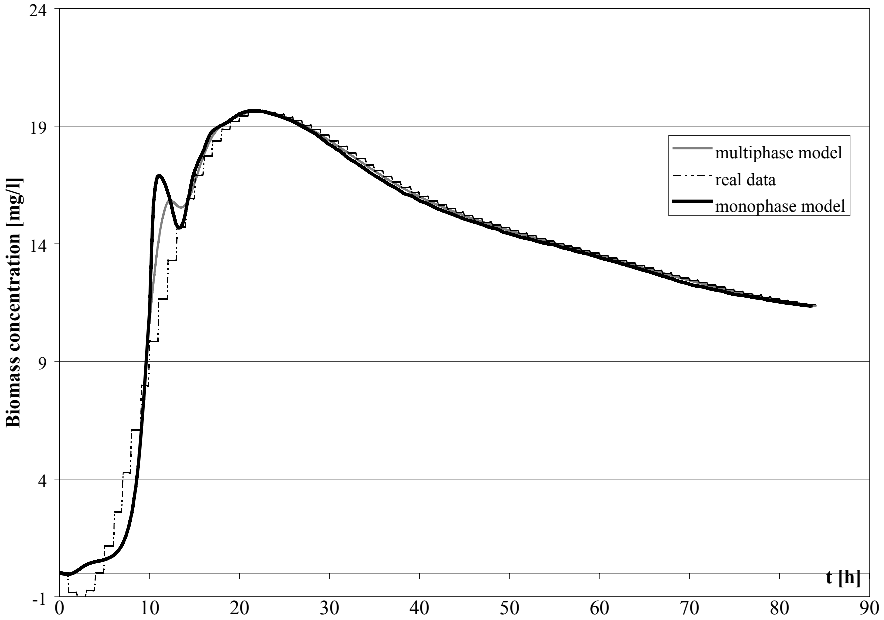

Fig. 1 – Biomass Concentration

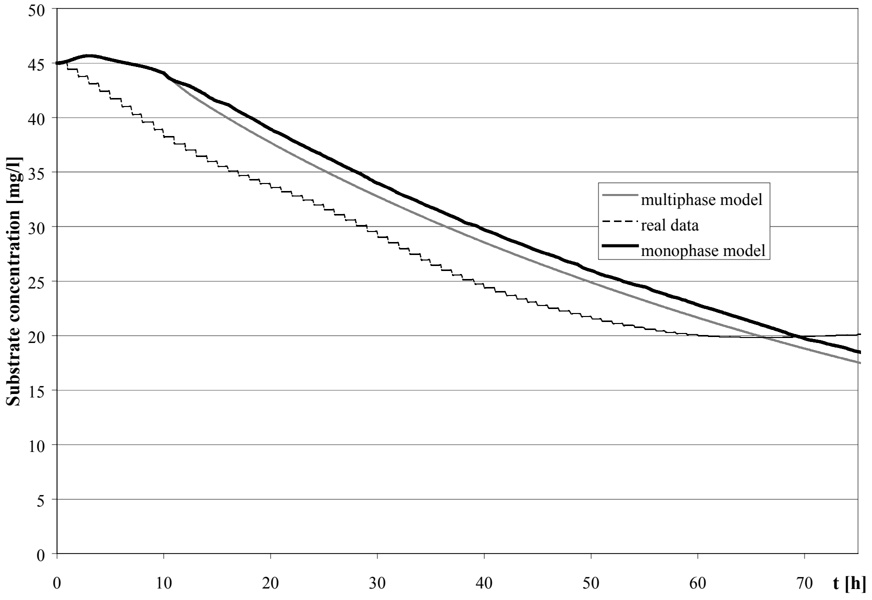

Fig. 2 - Substrate Concentration

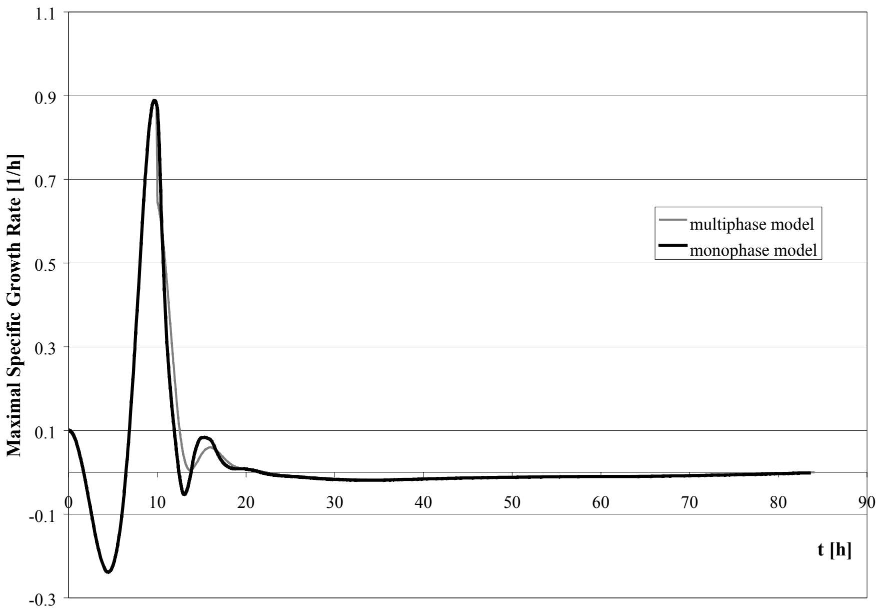

Fig. 3 - Maximal Specific Growth Rate

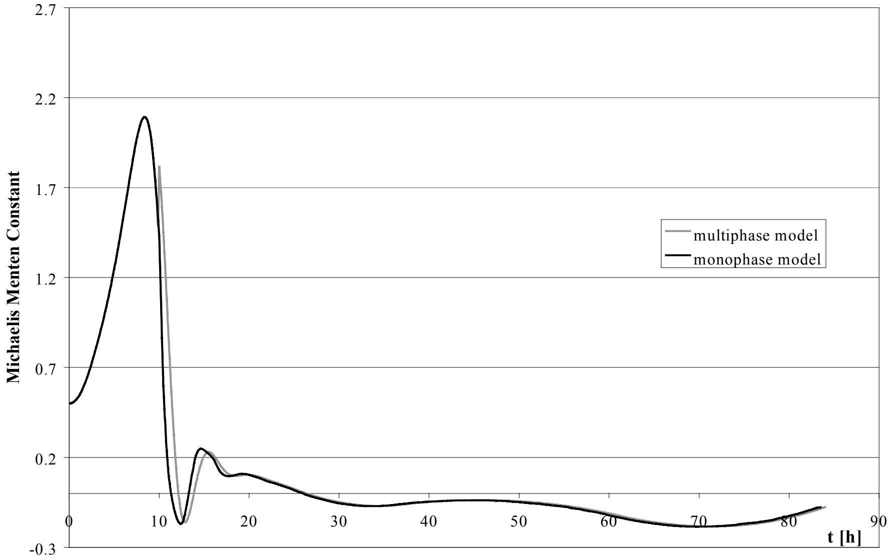

Fig. 4 – Michaelis Menten Constant

Table 1. Accuracy estimation of obtained models

Process phases	Average residual (SSR criterion)	
	Mono model	Multi model
Exponential phase	0.202819	0.142642
Lag phase	0.010124	0.007471
Global model	0.04344	0.030604

4. CONCLUSIONS

There are many reports about recursive techniques, which are designed for online use and take into account the characteristics of measurement errors.

The biological systems are characterized by complex, non-linear relationships involving poorly identified parameters. This makes them likely candidates for applying adaptive algorithms. To avoid explicitly modeling the specific rates, these can be treated as parameters, and estimated along with the state variables through the usage of Kalman filter [2,11,12]. The Kalman filter provides an estimate of the state that tries to minimize the inconistencies with all pieces of information in the least square sense. Presented technique for on-line state estimation of biomass is a software sensor for biomass inferential measurement, which is applicable on different biotechnological processes [2, 9,10].

The results show that the estimated values (with using a single and multi phase model) of biomass and product concentrations tend to experimental data with a satisfactory accuracy (Fig.1, 2). Modeling with multimodel approach observed a high accuracy in the time interval after the 15-th hour and a higher accuracy in the в transition between two phases – from lag to exponential, which is due to the chosen model structure. The results are sufficiently good and satisfy the requirements for model accuracy (fig.1,2) and gives less residual error (table 1).

Our expectations are, that at more detailed analysis of model structure, taking into account more growth factors, for example pH, which have influence on biomass synthesis, will be obtained a better accuracy of assessment. A higher accuracy of the obtained model would be achieved and with dividing a process to more than two phases, as well as if are used other modeling techniques – fuzzy logic, NN, under which is expected to achieve more smoothly transition between the process phases.

ACKNOWLEDGEMENT

This work is partially supported by the Project I 1406/2004 financed by the National Science Foundation with the Bulgarian Ministry of Education and Science.

5. REFERENCES

[1] Neeleman R., A.J.B. van Boxtel, *Estimation of specific growth rate from cell density measurements*,. Bioprocess and Biosystems Engineering, 24(3), 179-185, 2001.

[2] Vassileva S., Popova S., K.Tsekova (2005). *On-Line Parameters Estimation of the Model For Glucoamylase Production Using Aspergillus Niger Cells*. Int. Conf. ETAI'2005, Ohrid, Macedonia (to appear), 2005.

[3] Anderson, B.D.O., J.L. Moore, *Optimal filtering*. Prentice Hall, New Jersey, 1979.

[4] Bastin G., D. Dochain, *On-line estimation and adaptive control of bioreactors*, Elsevier, Amsterdam-Oxford-New York-Tokyo, 1990.

[5] Dochain D., M. Perrier, *Dynamic modelling, analysis, monitoring and control design for nonlinear bioprocesses*, Advances in Biochemical Engineering, Biotechnology, 56, 147-197, 1997.

[6] Claes JE, Impe JFV, *On-line estimation of the specific growth rate based on viable biomass measurements: experimental validation*. Bioprocess Eng 21:389-395, 1999.

[7] Farza M, H Hammouri, S Othman, K Busawon, *Nonlinear observers for parameter estimation in bioprocesses*, Chem Eng Sci52:4251-4267, 1997.

[8] Gee DA, Ramirez WF *On-line state estimation and parameter identification for batch fermentation*. Biotechnol Prog12:132-140, 1996.

[9] Vassileva S., S. Popova, K.Tsekova (2005). On-Line State and Parameters Estimation Of The Model For Glucoamylase Production Using Aspergillus Niger Cells. Biotechnology and Biotechnol.Equipment, ^{1}3 (to appear).

[10] Vassileva S., S. Popova, B.Tzvetkova (2005). Software Sensor for Protein Content and Protein-Synthesizing Ability Estimation of Biotechnologcal Processes. Second Intern. Sc. Conf. "COMPUTER SCIENCE", Chalkidiki, Greece, 30th September – 2nd October 2005(to appear).