SOFTWARE SYSTEM FOR GENERATING, IDENTIFICATION AND SIMULATION OF THE BIOTECHNOLOGICAL PROCESSES MODELS

Sergiu CARAMAN, Laurentiu FRANGU, Maria TURTOI

Department of Automatic Control and Electronics,
"Dunarea de Jos" University, Domneasca 111, 6200 Galati, Romania
Phone/fax: +40 36 460 182, email: Sergiu.Caraman@ugal.ro

Abstract: The paper deals with the development of a software system (BIOCALC) for the modelling, identification and simulation of the biotechnological processes. BIOCALC is based on the algorithm developed by George Bastin and Daniel Dochain in (Bastin, et al., 1990), starting with the reaction scheme of the process. The BIOCALC software generates the differential equations of the process, identifies the model coefficients and allows the model simulation comparing the results to the experimental data.

Keywords: biotechnology, reaction scheme, model identification, numerical simulation

1. INTRODUCTION

Biotechnological processes are very complex, strongly nonlinear and, therefore, quite difficult to model. As there are not enough accurate, reliable and not too expensive sensors, it is practically impossible to perform on-line measurements of the state variables of the process. Many researchers applied classic methods to model all kind of biotechnological processes (continuous, discontinuous and fed-batch), such as:
1. mass balance equations
2. reaction schemes
3. zoning of the biotechnological processes evolution.

All these methods lead to analytical models (structural-functional models) which can be used in control techniques.

Taking into consideration the bioprocesses particularities, the modelling of the biotechnological processes can be assisted by computer. It is wellknown the software system CAMBIO (Computer Aided Modelling of BIOprocesses), developed by the Biosystem team from the automation laboratory of Grenoble, France. It is the result of a deep analysis on the bioprocesses modelling made by A. Cheruy, R. Montellano and M. Farza (Farza, 1992). The

software mentioned above is based on a functional scheme, which is an intermediary stage between a verbal description and a set of mathematical equations. The user must specify the type of the process variables (substrate, biomass, enzyme, activating or inhibiting agent, reactant or final product) and the reaction links between variables. CAMBIO generates the differential equations of the process model.

The present paper deals with a similar software system, namely BIOCALC. It is based on the modelling algorithm presented in (Bastin, et al., 1990) and has as the main objective to generate the set of differential equations of the process' model. BIOCALC also allows to identify and to simulate the model of the bioprocess, compared to the experimental data.
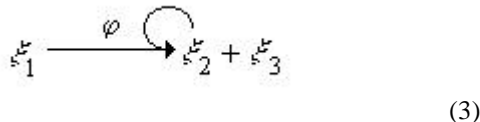
2. THE REACTION SCHEME OF A BIOPROCESS

A simple irreversible reaction involving two reactants and yielding a reaction product is commonly represented by a scheme of the following form:

$$x_1 + x_2 \xrightarrow{\ \ j\ \ } x_3 \qquad (1)$$

where x_1 and x_2 are two reactants which are irreversibly combined to give the reaction product x_3. The variable j is the reaction rate, i. e. the rate of consumption of the reactants, which is equal to the rate of formation of the product. Generally, the number of reaction components (reactants and products) is arbitrary and the scheme is written:

$$\sum x_i \xrightarrow{\ \ j\ \ } \sum x_j \qquad (2)$$

A catalyzed reaction is a reaction in which a component (namely catalyst) appears in both sides of the scheme. This means that the catalyst is assumed to be consumed and produced simultaneously at the same rate by the reaction (in such a way that the balance of this component is continuously in equilibrium). In an autocatalyst reaction, one product is a catalyst of its own production. In that case, the scheme is as follows:

$$\xi_1 \xrightarrow{\ \ \varphi\ \ } \circlearrowright \xi_2 + \xi_3$$

$$(3)$$

The feedback arrow indicates the presence of an autocatalyst x_2. It is a pseudoreactant which is not consumed by the reaction, but which can be accumulated in the reactor.

2. THE ALGORITHM FOR OBTAINING THE MODEL OF A BIOPROCESS

Once the reaction scheme of a biotechnological process is given, the derivation of a dynamical model is made fully systematic by applying the following rules (Bastin, et al., 1990):

R1) The reaction scheme of the process involves N components x_i (i=1,…,N) and M reactions (j=1,…,M). The reaction rates are denoted j_j (j=1,…,M).

R2) The dynamics of the concentration of each component x_i are written as follows:

$$\frac{dx_i}{dt} = \sum_{j \sim i} (\pm) k_{ij} j_j - D x_i - Q_i + F_i \qquad (4)$$

R3) The notation $j\sim i$ means that the summation is taken on the reactions with index j which involve the component with index i.

R4) k_{ij} are strictly positive constant yield coefficients, without dimension (i. e. units of mass / mass). They have a "-" sign when x_i is a reactant (i. e. it appears on the left-hand side of the reaction

scheme) and a "+" sign when x_i is a product of the reaction (i. e. it appears on the right hand side).

R5) Q_i is the rate of mass outflow of the component x_i from the reactor in the gaseous form.

R6) F_i is the mass feed rate in the reactor of the component x_i if it is an external substrate. Otherwise, F_i=0.

3. SOFTWARE IMPLEMENTATION OF THE BIOCALC SYSTEM

This section deals with the main characteristics of the BIOCALC system. For the prototype implementation it has been chosen the software environment Matlab, 5.3.0 version. The application contains a package of files of *m*-type and *mat*-type. The graphic interfaces were performed with *Graphical User Interface Tools* from Matlab. When the application is running there is not the possibility to fill the working memory by opening new windows. This advantage was obtained by an algorithm that suppresses the inactive windows using the windows of *modal*-type.

The first window (the welcome window) is presented in figure 1 and contains three main options: *modelling*, *identification* and *simulation*.
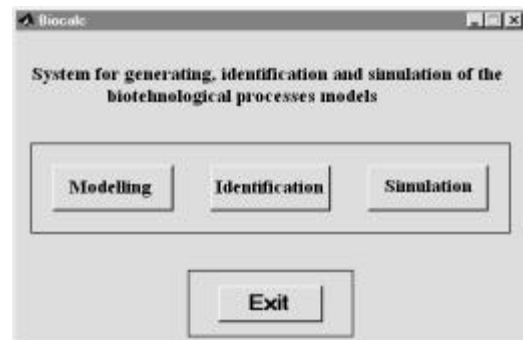

Fig. 1. The wellcome window

The *modelling* option allows obtaining the bioprocess model, the option *identification* calculates the coefficients of the model compared to the experimental data and the option *simulation* simulates the model and the experimental data.

Once the modelling option was selected, a new window with two sectors does appear on the screen (figure 2). In the left sector the process variables (reactants, reaction products and reaction rates) are introduced. In the right sector, the program builds the reaction scheme. The reactants and the reaction products are shown to the left side and to the right side respectively of the reaction rates list. The reaction is build selecting element by element the reactants from the first list, the reaction rate from the second list and the reaction products from the third

list, according to the algorithm presented in section 3 of the paper. Every reaction is save in a file and shown in the window under the three lists with process variables. The window from figure 2 also allows selecting the bioreactor type (continuous or discontinuous).
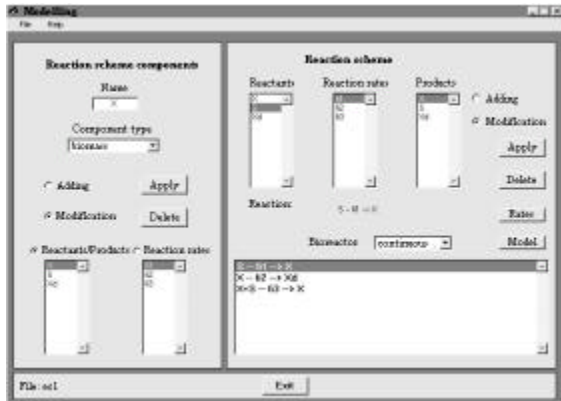

Fig. 2. The reaction scheme window

For the generating of the reaction scheme is necessary to define the reaction rate form (the *pushbutton Rate* in the window from figure 2). This will generate a new window (figure 3) with some possibilities to select the form of the reaction rate.
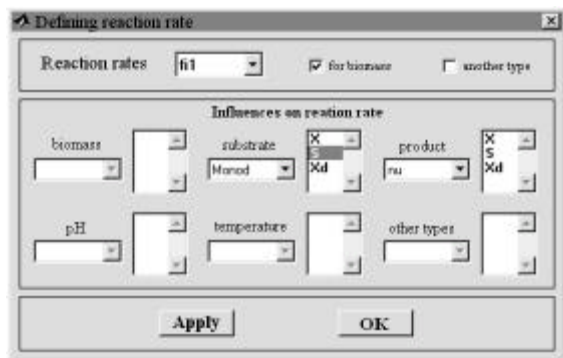

Fig. 3. The selection of the reaction rate

As one can see in figure 3, the software can model all the phases of a biotechnological process (the biomass developing, the decay phase, etc). BIOCALC allows to the user to express the reaction rate as function of biomass, substrate, product, pH, temperature and other variables. The user must select the specific reaction rate $m$as it can be seen in figure 4.

Fig. 4. The selection of the specific reaction rate type

The software also allows selecting a double dependence of the reaction rate, based on the multiplying principle. The reaction scheme can be obtained activating the *pushbutton* OK from the window in figure 2. The reactions generated by BIOCALC are presented in figure 5:

Fig. 5. The reaction scheme

The option *identification* allows the obtaining of the model coefficients based on experimental data. The experimental data are storred in a file and used by the identification algorithm (figure 6).
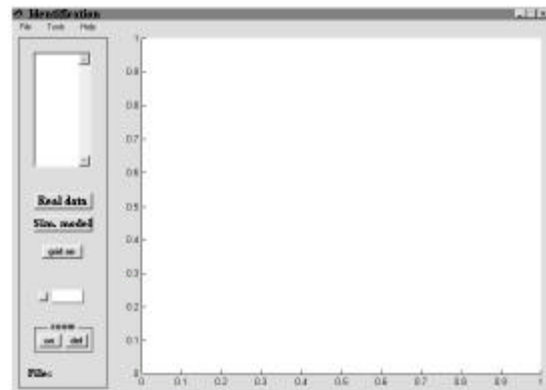

Fig. 6. The window for identification option

The window contains a *menubar* divided in three option groups: *File*, *Tools* and *Help* (figure 7).

Fig. 7. The *menubar* window

Very important are the options *Run* (begin the identification process) and *Clear* (delete the curves from the right side of the window). While the identification process is running, a *waitbar* window appears on the screen (figure 8).
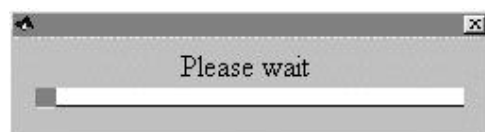

Fig. 8. The *waitbar* window

The calculus of the model parameters is made by a search procedure (unidimensional search) in the parameters space. The search procedure minimizes the distance between the model curves and experimental data. Before the identification process starts the initial values of the model coefficients must be introduced (figure 9).

Fig. 9. The initial values of the model coefficients are introduced

The *simulation* option calculates the reactants and reaction products concentrations as functions of time *t* for an identified set of the model coefficients or another set chosen by the user. The corresponding window is similar to the one used for the *identification* option.

4. CASE STUDY

Let's consider the biotechnological process given by the following reaction scheme:

$$S \xrightarrow{\ j_g\ } X \tag{5}$$

$$X \xrightarrow{\ j_d\ } X_d \tag{6}$$

$$S + X \xrightarrow{\ j_m\ } X \tag{7}$$

Based on the algorithm presented in section 3, BIOCALC determines the following model (figure 5):

$$\frac{dX}{dt} = k_{11} j_1 - k_{12} j_2 - DX \tag{8}$$

$$\frac{dS}{dt} = -k_{21} j_1 - k_{23} j_3 - DS + F_2 \tag{9}$$

$$\frac{dX_d}{dt} = k_{32} j_2 - DX_d \tag{10}$$

where

$$j_1 = m(S) \cdot X \tag{11}$$

$$j_2 = k_d \cdot X \quad cu \ k_d = const. \tag{12}$$

$$j_3 = k_m \cdot X \quad cu \ k_m = const. \tag{13}$$

The specific reaction rate is given by the Monod form.

$$m(S) = \frac{m^* S}{k_M + S} \tag{14}$$

Replacing relations (11) - (14) in the model equations (8) - (10), results

$$\frac{dX}{dt} = k_{11} \frac{m^* S}{k_M + S} X - k_{12} k_d - DX \tag{15}$$

$$\frac{dS}{dt} = -k_{21} \frac{m^* S}{k_M + S} X - k_{23} k_m - DS + F_2 \tag{16}$$

$$\frac{dX_d}{dt} = k_{32} k_d - DX_d \tag{17}$$

which can be rewritten in the following form:

$$\frac{dX}{dt} = k_X \frac{S}{k_M + S} X - k_{12} k_d - DX \tag{18}$$

$$\frac{dS}{dt} = -k_S \frac{S}{k_M + S} X - k_{23} k_m - DS + F_2 \tag{19}$$

$$\frac{dX_d}{dt} = k_{32} k_d - DX_d \tag{20}$$

with

$$k_X = k_{11} m^* \tag{21}$$

$$k_S = k_{21} m^* \tag{22}$$

The coefficients that must be identified are: k_X, k_S, $k_{12}, k_{23}, k_{32}, k_M$.

It has been considered the following values:

k_m=0.5
k_d=0.5
F_2=25
D=0.3

The initial values of the unknown coefficients are set to 1. The integrating parameters have the following values:

t_0=0   (the initial time for integrating)
t_f=10h,   (the final time for integrating)
tol=10^{-5}, (the tolerance)
$step$=0.05, (the initial search step = 5% of
                 the initial value of the coefficient)

The initial values of the state variables:

X_0=1 [g/l]
S_0=5 [g/l]
X_{d0}=0.1 [g/l]

The experimental data used for the model identification are presented in figure 10.
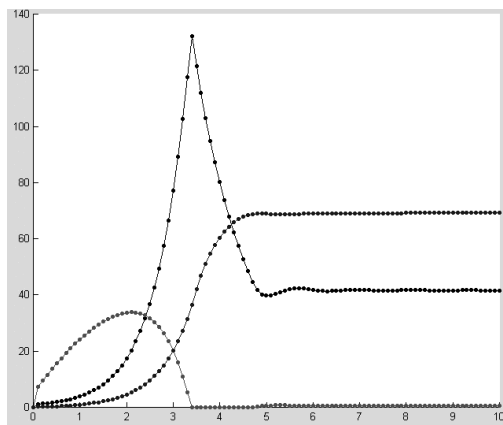

Fig. 10. The set of experimental data

It has been used an extended set of experimental data for the identification procedure, obtained by interpolation.

The final values of the unknown coefficients determined by BIOCALC system are the following:

k_X=2.47
k_S=0.3324
k_{12}=0.99274
k_{23}=0.9863
k_{32}=1.0168
k_M=1.0006

The results of the model simulation with the set of coefficients identified by BIOCALC system are presented in figure 11.
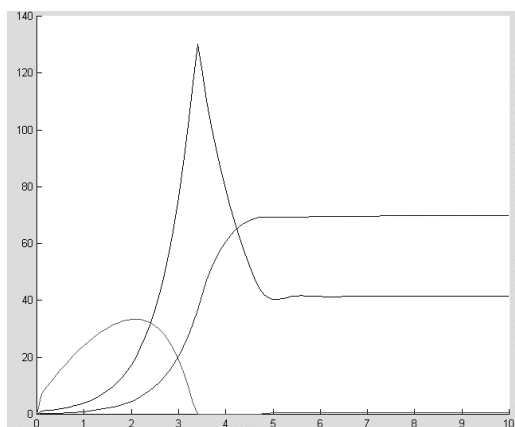

Fig.11. The results of the model simulation with the set of coefficients identified by BIOCALC system

5. CONCLUSIONS

The following conclusions can be drawn:

1. BIOCALC system is an automatic generator of the models, which can be implemented on a computer, connected "on-line" to the bioreactor.
2. The functions implemented so far are limited mainly due to the biotechnological processes complexity and the time variations of the parameters. They influence very strong the process evolution. The large number of variants for parametrizing the most important coefficients of the biotechnological processes also influence the process.
3. The next step is to add new variants for parametrizing the specific growth rate as well as to diversify the identification and the integration methods for the model of the biotechnological process.
4. The BIOCALC prototype was implemented using the MATLAB environment due to its abilities to integrate and represent graphics. The final version will be implemented using VISUAL C++.

REFERENCES

Bastin, G. and D. Dochain (1990). *On-line Estimation and Adaptive Control of Bioreactors*, pp. 375, Elsevier.

Farza, M. (1992). *Récents développements dans l'aide à la modélisation et à l'estimation des bioprocédés*, ph. D Thesis INPG Laboratoire d'Automatique de Grenoble, France.

Bastin, G. and Dochain, D., (1986). On-line Estimation of Microbial Specific Growth Rates, Automatica, Vol. 22, No. 6, Pergamon Journals Ltd., pp. 705-709.