Application of Genetic Algorithm to Tuning a PID Controller for Glucose Concentration Control

TSONIO SLAVOV[1], OLYMPIA ROEVA[2]
[1] Department of Automatics and Systems Engineering
Technical University of Sofia
8 Kliment Ohridski Blvd., 1756 Sofia,
BULGARIA
ts_slavov@tu-sofia.bg

[2] Department of Bioinformatics and Mathematical Modeling
Institute of Biophysics and Biomedical Engineering
Bulgarian Academy of Sciences
105 Acad G. Bonchev Str., 1113 Sofia,
BULGARIA
olympia@biomed.bas.bg

Abstract: - The paper presents a feedforward feedback (PID) controller designed for control of glucose concentration during the *E. coli* fed-batch cultivation process. The controller is used to control the feed rate and to maintain glucose concentration at a desired set point. Taking into account the measurement system particularities, the modified process model is proposed. An equation for correction of the measured glucose based on Kalman filter estimates of biomass concentration and bacteria growth rate is suggested. To achieve good closed-loop system performance genetic algorithm tuning of the PID controller is used. As a result, the optimal PID controller settings are obtained. For a short time the controller sets the control variable and maintains it at the desired set point during the process. Based on the proposed model correction, the estimations of the process parameters are brought closer to the real values. Tuning of the controller on the basis of a genetic algorithm leads to higher level of accuracy and efficiency of the system performance.

Key-Words: - *E. coli* Cultivation, PID Controller, Tuning, Genetic Algorithm.

1 Introduction

Fed-batch cultivation process has been widely employed for the production of various bioproducts including primary and secondary metabolites, proteins, and other biopolymers. Fed-batch culture is especially beneficial when changing nutrient concentrations affect the productivity and yield of a desired cultivation product. Since both overfeeding and underfeeding of nutrient is detrimental to cell growth and product formation, development of a suitable feeding control strategy is critical in fed-batch cultivation.

Considering cultivation processes, different control loops are implemented for control of temperature, pH, and dissolved oxygen, as well as for the control of volume and anti foam. However, commercially available controllers are applied only to such well established measurement systems. For example, there is a shortage of glucose concentration control systems. The main reason for that is related to the difficulties in fast and reliable on-line measurement of the process variables, especially those for the substrate concentration.

The most widely used controller for industrial cultivation processes is the proportional-integral-derivative (PID) controller. P-term reduces the error but does not eliminate it, i.e. an offset between the actual and desired value will normally exist. The additional I-term corrects the error that occurs between the desired value (set point) and the process output. Inclusion of the I-term makes the control system more likely to oscillate. Inclusion of the D-term improves the speed of the responses, and consequently serves to suppress the influence of the disturbance more strongly [1-2]. However, the D-term functions are effective only when the controller parameters are appropriately tuned. Controller tuning is a subjective procedure and it is certainly process dependent. Usually, PID controllers are poorly tuned due to highly changing dynamics of most bioprocesses, caused by the non-linear growth of the cells and the changes in the overall metabolism. Tuning a PID controller appears to be

conceptually intuitive but can be rather hard in practice if cultivation processes are considered. Due to changes of the system parameters, conventional PID controllers result in sub-optimal corrective actions and hence require retuning. Therefore, a higher degree of experience and technology is required for controller tuning in a real plant. While for control of continuous cultivation processes the controller tuning could be done with traditional methodology [3], for fed-batch cultivation processes such methodologies are inapplicable.

In this case, as an alternative of the quality controller tuning, metaheuristics could be applied, since the tuning procedure is a big challenge for the conventional optimization methods. Heuristics can obtain suboptimal solution in ordinary situations and optimal solution in particular ones. Since the considered problem has been known to be NP-complete, the use of heuristic techniques can solve this problem more efficiently. In the probabilistic optimization group, genetic algorithms (GA)-based methods are considerable part, which extensively are proposed in the literature. GA originated from the studies of cellular automata, conducted by John Holland and his colleagues at the University of Michigan. Holland's book [4], published in 1975, is generally acknowledged as the beginning of the research of genetic algorithms. The GA is a model of machine learning which derives its behaviour from a metaphor of the processes of evolution in nature [5]. This is done by the creation within a machine of a population of individuals represented by chromosomes. A chromosome could be an array of real numbers, a binary string, a list of components in a database, all depending on the specific problem. The GA are highly relevant for industrial applications, because they are capable of handling problems with non-linear constraints, multiple objectives, and dynamic components – properties that frequently appear in real-world problems [2, 5]. Since their introduction and subsequent popularization [4], GA have been frequently used as an alternative optimization tool to the conventional methods [5, 6] and have been successfully applied to a variety of areas, and still find increasing acceptance [7-10].

Authors in [3] have applied a genetic algorithm to tuning of a PID controller for a continuous model bioreactor. Compared to some traditional tuning methods, the obtained results reflect that the use of GA based controllers improves the performance of the process in terms of time domain specifications, set point tracking, and regulatory changes. Authors in [11, 12] show that applying GA global and local optimal solution can simultaneously be achieved and the most appropriate parameter of PID controller can be selected for the given plant and system during operation. In [13], much more improved performance of considered GA tuned controller than the conventional ones has been revealed in terms of overshoot, settling time, etc. The generality of GA combined with its intuitiveness, fast convergence, modest processing requirements and most importantly minimal system specific information result in increased use of this technique for tuning of PID controllers [14-17].

In this paper, a glucose concentration feed forward feedback control system, in the presence of measurement and process noises, is designed. *E. coli* MC4110 fed-batch cultivation process is considered as a case study. The feedback part of control signal is based on the GA tuned universal digital PID controller. The main contribution of the paper is the proposed correction of the glucose measurement system time delay. This correction is formed from the process variables estimates evaluated by the designed extended Kalman filter (EKF). In this way, the influence of measurement time delay on control system performance is significantly reduced. The designed control system successfully keeps the glucose concentration at a desired low level. Thus, a prevention of the growth inhibition, based on (in a result of) substrate excess and optimal operational regime for cultivation process are achieved.

2 *E. coli* MC4110 Fed-Batch Cultivation Model

Cultivation of recombinant micro-organisms, e.g. *E. coli*, in many cases is the only economical way to produce pharmaceutical bio-chemicals such as interleukins, insulin, interferons, enzymes and growth factors. *E. coli* is still the most important host organism for recombinant protein production.

For *E. coli* MC4110 cultivation model identification, real experimental data are used [18, 19]. They were obtained in the Institute of Technical Chemistry, Hannover University. The cultivation conditions, data measurements and GA identification procedure are discussed in [18]. As a result, on the basis of feeding rate data and off-line measurements of biomass and on-line data of substrate (glucose) measurements, a mathematical model is obtained [18]. It could be represented as follows:

$$\begin{cases} \dot{\mathbf{X}}(\mathbf{t}) = \mathbf{f}(\mathbf{X},\mathbf{Q}) + \boldsymbol{\eta}(\mathbf{t}) \\ \gamma_S = \mathbf{H}\mathbf{X} + \xi(t) \end{cases}, \quad (1)$$

$$\mathbf{X}(\mathbf{t}) = \left[\gamma_X(t) \quad \gamma_S(t) \quad V(t) \quad \mu_{\max}(t)\right]^{\mathrm{T}} \tag{2}$$

$$\mathbf{f}(\mathbf{X}, Q) = \begin{bmatrix} \mu_{\max}(t)\dfrac{\gamma_S(t)}{k_S+\gamma_S(t)}\gamma_X(t) - \dfrac{Q(t)}{V(t)}\gamma_X(t) \\ -\dfrac{1}{Y_{S/X}}\mu_{\max}(t)\dfrac{\gamma_S(t)}{k_S+\gamma_S(t)}\gamma_X(t) + \dfrac{Q(t)}{V(t)}\left(\gamma_{in}-\gamma_S(t)\right) \\ Q(t) \\ 0 \end{bmatrix} \tag{3}$$

$$\mathbf{H} = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \tag{4}$$

$$\boldsymbol{\eta}(\mathbf{t}) = \left[\eta_{\gamma_X}(t) \quad \eta_{\gamma_S}(t) \quad \eta_V(t) \quad \eta_{\mu_{\max}}(t)\right]^{\mathrm{T}} \tag{5}$$

Numerical values of the parameters and initial conditions are according to [18]:

$k_S = 0.012$ g l^{-1},
$Y_{S/X} = 0.5$,
$\gamma_{in} = 100$ g l^{-1},
$t_0 = 6.68$ h,
$\gamma_X(t_0) = 1.25$ g l^{-1} and $\gamma_S(t_0) = 0.8$ g l^{-1}.

In addition, the process noise $\boldsymbol{\eta}(\mathbf{t})$ and measurement noise $\xi(t)$ are set to zero mean white Gaussian noises. The corresponding variances are:

$D_{\eta_{\gamma X}} = 0.001$ g^2 l^{-2} h^{-1},
$D_{\eta_{\gamma S}} = 0.001$ g^2 l^{-2} h^{-1}, $D_{\eta_V} = 0$ l^2 h^{-3},
$D_{\eta_{\mu_{\max}}} = 0.05$ l h^{-3} and $D_{\xi} = 0.0025$ g^2 l^{-2} h^{-1}.

The samples containing the cells are pumped out of the bioreactor for the glucose measurement. During the analysis time, the cells still consume glucose. In the beginning of the cultivation, at lower biomass concentration, consumed glucose is negligible, but after that the biomass concentration has grown and the consumed glucose is significant. As a result, the measured glucose concentration is not accurate (Fig. 1). To overcome this problem, some modifications of the process model are required.

In [19], a correction of glucose concentration measurement using constant average specific growth rate is considered. The proposed correction of glucose concentration through variable specific growth rate (i.e. variable maximum specific growth rate) is used to ensure more closely the real dynamics of the cultivation process conditions. The proposed equation for glucose concentration measurement correction has the following form:

$$\gamma_{S_{COR}}(t) = \gamma_S(t) + \frac{\mu(t)\gamma_X(t)}{Y_{X/S}}\Delta t \tag{6}$$

The time delay is set to $\Delta t = 60$ s. The specific grow rate $\mu(t)$ is described by Monod kinetics as:

$$\mu(t) = \mu_{\max}(t)\frac{\gamma_S(t)}{k_S+\gamma_S(t)}. \tag{7}$$

As a result, the corrected glucose measurements are closer to the real one.

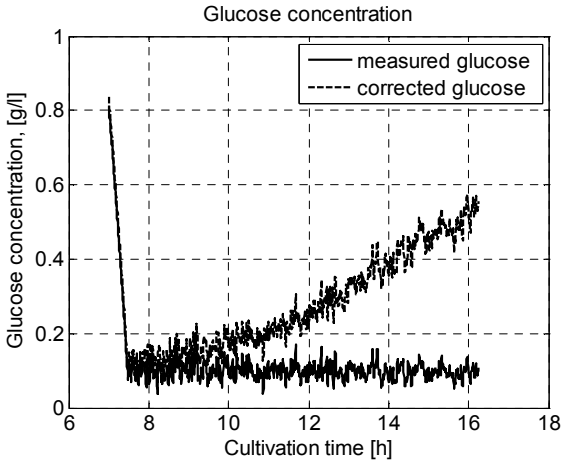

Fig. 1. Measured glucose concentration and the corrected one

3 Background of the Control Algorithm

The structure of the herewith designed control system is shown in Fig. 2. The presented feedback control algorithm is described based on [20-22], as follows:

$$\begin{aligned} u_{fb}(s) = {} & K_p\left(br(s)-\gamma_S(s)\right) + \\ & + \frac{K_p}{T_i}\left(r(s)-\gamma_S(s)\right) + \\ & + \frac{T_d s}{1+\dfrac{T_d s}{N}}\left(cr(s)-\gamma_S(s)\right) \end{aligned} \tag{8}$$

The control variable $u_{fb}(s)$ is a sum of three terms: P-term which is proportional to the control error (difference between set point (reference signal) $r(s)$ and the process output $\gamma_S(s)$), I-term which is proportional to the integral of the control error and, D-term which is proportional to the derivative of this error. The controller parameters are K_p, T_i, T_d, b, c and $\frac{T_d}{N}$. To reduce the influence of measurement noise, a first-order low pass filter is used. The coefficients b and c are used to weigh out $r(s)$ respectively in P-term and in D-term of the controller. Typically, the coefficients are chosen to be $0 \le b \le 1$, $0 \le c \le 1$. In industrial applications, b and c are chosen to be equal to 0 or 1 [20, 21].

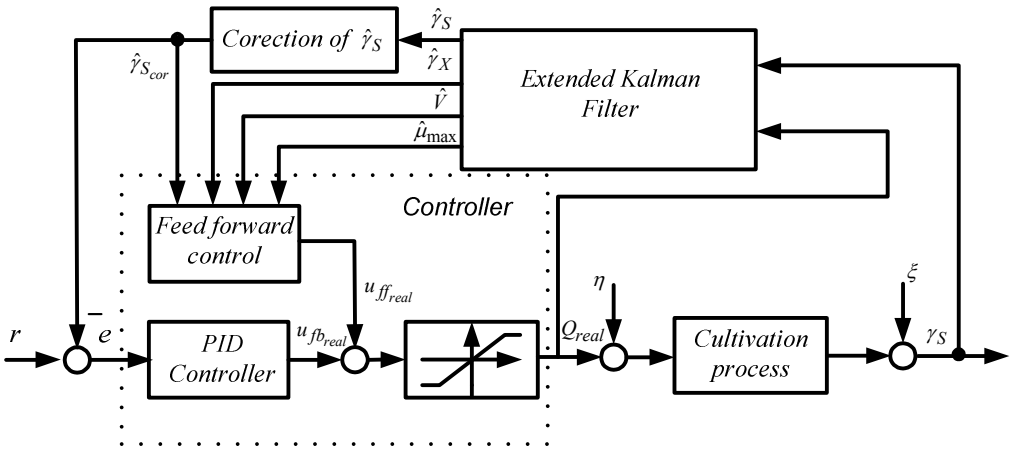

Fig. 2. Structure of the designed control system

The considered here cultivation process is characterized with exponentially rising load disturbance (the biomass concentration γ_X is exponentially rising during cultivation). Due to this particularity it is proposed the coefficients b and c to take values greater than 1. Considering real applications usually digital PID controller is implemented. There are many techniques for discretization [23]. Here for discretization of the PID controller (assumption (8)) backward Euler method [24] is used. The mathematical description of the designed digital PID controller is:

$$u_{fb}(k) = u_p(k) + u_i(k) + u_d(k) \tag{9}$$

$$u_p(k) = K_p\left(br(k) - \gamma_S(k)\right) \tag{10}$$

$$\begin{aligned} u_i(k) &= u_i(k-1) + \\ &+ b_{i1}\left(r(k) - \gamma_S(k)\right) + \\ &+ b_{i2}\left(r(k-1) - \gamma_S(k-1)\right) \end{aligned} \tag{11}$$

$$\begin{aligned} u_d(k) &= a_d u_d(k-1) + \\ &+ b_d\left(cr(k) - cr(k-1) - \gamma_S(k) + \gamma_S(k-1)\right) \end{aligned} \tag{12}$$

where:

$$\begin{aligned} b_{i1} &= K_p\frac{T_0}{T_i}, \\ b_{i2} &= 0, \\ a_d &= \frac{Td}{Td + NT_0}, \\ b_d &= K_p\frac{T_d N}{Td + NT_0}. \end{aligned} \tag{13}$$

The control variable used to control the feed rate has the following form:

$$Q(k) = u_{fb}(k) + u_{ff}(k) \tag{14}$$

where

$$u_{ff}(k) = \frac{1}{Y_{S/X}}\frac{V(k)\mu(k)\gamma_X(k)}{\gamma_{in} - \gamma_S} \tag{15}$$

is feedforward term obtained from γ_S steady state conditions.

For calculation of the control variable assumption (14) on-line measurements of $\gamma_X(k)$ and V(k) are required. Due to the lack of such data an EKF is designed. Based on discretization of process model (assumptions (1) - (5)) the following EKF is obtained:

$$\left|\begin{aligned} &\hat{\mathbf{X}}(k+1) = \mathbf{f_d}\left(\hat{\mathbf{X}}(k)\right) + \\ &+\mathbf{K_{EKF}}(k+1)\left(\gamma_S(k+1) - \hat{\gamma}_S(k+1)\right), \\ &\hat{\gamma}_S(k+1) = \mathbf{H}\hat{\mathbf{X}}(k+1), \end{aligned}\right. \tag{16}$$

$$\hat{\mathbf{X}}(0) = \begin{bmatrix}1.25 & 0.8 & 1.35 & 0.55\end{bmatrix}^{\mathrm{T}},$$

$$\mathbf{f_d}\left(\hat{\mathbf{X}}(k)\right) = \hat{\mathbf{X}}(k) + T_0\mathbf{f}\left(\hat{\mathbf{X}}(k)\right) \tag{17}$$

where:

$\hat{\mathbf{X}}(\cdot)$ – the estimate of $\mathbf{X}(\cdot)$, g·l^{-1};

$\hat{\gamma}_S(\cdot)$ – the estimate of glucose concentration $\gamma_S(\cdot)$, g·l^{-1};

$\mathbf{K_{EKF}}(\cdot)$ – the EKF gain, -.

The EKF gain is presented as:

$$\begin{aligned} &\mathbf{K_{EKF}}(k+1) = \\ &= \left[\left(\mathbf{F}(k)\mathbf{P}(k)\mathbf{F}(k)^{\mathrm{T}} + \mathbf{D_{\eta d}}\right)H^{\mathrm{T}}\right]\times \\ &\times\left[H\left(\mathbf{F}(k)\mathbf{P}(k)\mathbf{F}(k)^{T} + \mathbf{D_{\eta d}}\right)H^{\mathrm{T}} + D_{\xi}\right]^{-1} \end{aligned} \tag{18}$$

The noise covariance matrixes have the following forms:

$$\mathbf{D}_{\boldsymbol{\eta}} = T_0^2 \begin{bmatrix} 0.001 & 0 & 0 & 0 \\ 0 & 0.001 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.05 \end{bmatrix}, \quad (19)$$

$D_{\xi} = 0.0025,$

where: $\mathbf{D}_{\boldsymbol{\eta d}}$ is a covariance matrix of discrete-time process noise $\boldsymbol{\eta}_{\mathbf{d}}(k) = T_0 \boldsymbol{\eta}(t)$.

The $\mathbf{P}(\cdot)$ and $\mathbf{F}(k)$ are obtained from:

$$\mathbf{P}(k+1) = \left(\mathbf{I} - \mathbf{K}_{\mathbf{EKF}}(k+1) H^{\mathrm{T}}\right) \times$$

$$\times \left(\mathbf{F}(k)\mathbf{P}(k)\mathbf{F}(k)^{\mathrm{T}} + \mathbf{D}_{\boldsymbol{\eta d}}\right), \quad (20)$$

$$\mathbf{P}(0) = \mathrm{diag}(0.02 \quad 0.02 \quad 0 \quad 10)$$

$$\mathbf{F}\left(\hat{\mathbf{X}}(k)\right) = \mathbf{I}_{\mathbf{4}} + T_0 \boldsymbol{\Phi}\left(\hat{\mathbf{X}}(k)\right), \quad (21)$$

where

$$\boldsymbol{\Phi}(\hat{\mathbf{X}}(k)) = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (22)$$

$$a_{11} = \frac{\hat{\mu}_{\max}(k)\hat{\gamma}_S(k)}{K_{\gamma_S} + \hat{\gamma}_S(k)} - \frac{Q(k)}{\hat{V}(k)},$$

$$a_{12} = \frac{\hat{\gamma}_X(k)\hat{\mu}_{\max}(k)\left(K_{\gamma_S} + \hat{\gamma}_S(k)\right)}{\left(K_{\gamma_S} + \hat{\gamma}_S(k)\right)^2} - \frac{\hat{\gamma}_X(k)\hat{\mu}_{\max}(k)\hat{\gamma}_S(k)}{\left(K_{\gamma_S} + \hat{\gamma}_S(k)\right)^2},$$

$$a_{13} = \frac{Q(k)\hat{\gamma}_X(k)}{\hat{V}^2(k)}, \quad a_{14} = \frac{\hat{\gamma}_X(k)\hat{\gamma}_S(k)}{K_{\gamma_S} + \hat{\gamma}_S(k)},$$

$$a_{21} = -\frac{\hat{\mu}_{\max}(k)\hat{\gamma}_S(k)}{Y_{S/X}\left(K_{\gamma_S} + \hat{\gamma}_S(k)\right)},$$

$$a_{22} = -\frac{Q(k)}{\hat{V}(k)} + \frac{\hat{\gamma}_X(k)\hat{\mu}_{\max}(k)\hat{\gamma}_S(k)}{Y_{S/X}\left(K_{\gamma_S} + \hat{\gamma}_S(k)\right)^2} - \frac{\hat{\gamma}_X(k)\hat{\mu}_{\max}(k)\left(K_{\gamma_S} + \hat{\gamma}_S(k)\right)}{Y_{S/X}\left(K_{\gamma_S} + \hat{\gamma}_S(k)\right)^2},$$

$$a_{23} = -\frac{Q(k)\left(\gamma_{in} - \hat{\gamma}_S(k)\right)}{\hat{V}^2(k)},$$

$$a_{24} = \frac{\hat{\gamma}_X(k)\hat{\gamma}_S(k)}{Y_{S/X}\left(K_{\gamma_S} + \hat{\gamma}_S(k)\right)},$$

Finally, the real control variable has the following form:

$$Q_{real}(k) = u_{fb_{real}}(k) + u_{ff_{real}}(k), \quad (23)$$

where

$$u_{fb_{real}}(k) = u_{p_{real}}(k) + u_{i_{real}}(k) + u_{d_{real}}(k),$$

$$u_{ff_{real}}(k) = \frac{1}{Y_{S/X}} \frac{\hat{V}(k)\hat{\mu}(k)\hat{\gamma}_X(k)}{\gamma_{in} - \hat{\gamma}_{S_{cor}}}, \quad (24)$$

$$u_{p_{real}}(k) = K_p\left(br(k) - \hat{\gamma}_{S_{cor}}(k)\right), \quad (25)$$

$$u_{i_{real}}(k) = u_{i_{real}}(k-1) + b_{i1}\left(r(k) - \hat{\gamma}_{S_{cor}}(k)\right) + b_{i2}\left(r(k-1) - \hat{\gamma}_{S_{cor}}(k-1)\right), \quad (26)$$

$$u_{d_{real}}(k) = a_d u_{d_{real}}(k-1) + b_d\left(cr(k) - cr(k-1) - \hat{\gamma}_{S_{cor}}(k) + \hat{\gamma}_{S_{cor}}(k-1)\right). \quad (27)$$

The designed control system variable is based on EKF estimations of biomass concentration, corrected glucose concentration, specific growth rate and volume. The estimate of the corrected glucose concentration measurements $\hat{\gamma}_{S_{cor}}(k)$ are obtained as:

$$\hat{\gamma}_{S_{cor}}(k) = \hat{\gamma}_S(k) + \frac{\hat{\mu}(k)\hat{\gamma}_X(k)}{Y_{X/S}}\Delta t, \quad (28)$$

where

$$\hat{\mu}(k) = \hat{\mu}_{\max}(k)\frac{\hat{\gamma}_S(k)}{K_{\gamma_S} + \hat{\gamma}_S(k)}. \quad (29)$$

To provide control action designed for specific process requirements, tuning of the PID controller parameters (K_p, T_i, T_d, b, c and N) is required. With respect to control systems based on linear plant models there are many classical and novel or modified approaches for PID controller parameters tuning [20, 21, 25, 26]. These methods are inapplicable to the non-linear control system considered here. The regarded fed-batch cultivation process cannot be linearized around an equilibrium point. If a linear approximation is found, the resulting model will be valid only for a small range around the linearization point. The controller tuned by the linear model will work properly only for this limited range and for a very small time interval. Therefore, to achieve the best overall PID control it is necessary to use non-classical tuning methods for the entire operating envelope of the given system. In this work for PID controller parameters tuning, based on a control system (Fig. 2), GA are applied.

4 Background of the Genetic Algorithm for PID Controller Tuning

Outline of the applied here GA could be presented as:

1. **[Start]** Generate random population of n chromosomes.
2. **[Fitness]** Evaluate the fitness $g(x)$ of each chromosome x in the population.
3. **[New population]** Create a new population by repeating the following steps until the new population is complete.
   - 3.1. **[Selection]** Select two parent chromosomes from a population according to their fitness.
   - 3.2. **[Crossover]** With a crossover probability, cross over the parents to form new offspring.
   - 3.3. **[Mutation]** With a mutation probability, mutate new offspring at each locus.
   - 3.4. **[Accepting]** Place new offspring in the new population.
4. **[Replace]** Use the newly generated population for a further run of the algorithm.
5. **[Test]** If the end condition is satisfied, stop and return the best solution in current population.
6. **[Loop]** Go to step **2**.

Each individual represents a possible solution, and a set of individuals form a population. In a population, the fittest individuals are selected for mating. The individuals in the population go through a process of evolution which is, according to Darwin, based on the principles of mutation and selection; however, the modern biological evolution theory includes also crossover and isolation mechanisms improving the adaptiveness of the living organisms to their environment. With GA, elements are swapped between individuals as if by sexual combination and reproduction (crossover), and others are changed at random (mutation). New generations appear from clones of the current population, in proportion to their fitness: a single objective function of the parameters that returns a numerical value to distinguish between good and bad solutions. Fitness is then used to apply selection pressure to the population in a 'Darwinian' fashion (survival of the fittest) [5].

The parameters of a GA significantly affect the speed of convergence near the optimal solution, as well as the accuracy of the solution itself. A brief description of the GA parameters used here is presented below. The parameters choice is made on the basis of previous simulation tests [27, 28].

Encoding

A binary 20-bit encoding is considered. Binary representation is the most common one, mainly because of its relative simplicity. The best known selection mechanism, namely, roulette wheel selection, is used in the proposed GA.

Genetic operators

The genetic operators used in this GA are reproduction, crossover and mutation. Offspring are normally different from parents due to the process of genetic information exchange, e.g. chromosome crossover. However, in GA, the reproduction process is merely a simple coping action which passes the parent's genetic information to the offspring. The reproduction process usually acts as a complementary process of crossover and the offspring are either created by reproduction or crossover. Crossover is an extremely important component of GA, as it is responsible for searching through the solution space. Crossover can be quite complicated and depends (as well as the technique of mutation) mainly on the encoding of chromosomes. Here, double point crossover is employed. After a crossover is performed, mutation takes place. Mutation reintroduces diversity into the population. In accepted encoding here a bit inversion mutation is used. This prevents the solution from converging to some local optimal solutions; thereby the global optimal solution can be obtained.

Genetic parameters

Some particularly important parameters of GA are the population size (number of individuals) and number of generations. If the number of chromosomes is too small, GA has fewer possibilities to perform crossover and only a small part of search space is explored. Small populations cause premature convergence. Large population size takes longer time to run but converges faster than smaller populations. A number of 100 individuals is used as a trade-off. The higher crossover rates lead to better relative fitness function values. A crossover rate of 0.70 was used. Zero mutation probability causes premature convergence to a suboptimal value. Higher mutation probabilities, on the other hand, cause fluctuations and disturb convergence. Generally, the mutation probability should be high enough to expand the search space, but low enough to bring about convergence. A mutation probability of 0.05% was used as a trade-off. The chosen GA operators and parameters are summarized in Table 1.

Table 1. Genetic algorithm operators and parameters

Operator	Type	Parameter	Value
encoding	binary	crossover rate	0.70
crossover	double point	mutation rate	0.05
mutation	bit inversion	encoding precision	20
selection	roulette wheel selection	number of individuals	50
fitness function	linear ranking	number of generations	100

The substantial points of genetic algorithm for PID controller parameters tuning are initialization of algorithm parameters, representation of chromosomes, as well as choice of objective function. Some brief discussion about these points is presented below.

Initialization of algorithm parameters

The most appropriate GA parameters and operators, based on previous authors' results on the effects of the different GA parameters on the outcome of the GA [18, 27, 28] are used.

Representation of chromosomes

Representation of chromosomes is a critical part of the GA application. In order to use GA to identify controller parameters, it is necessary to encode the parameters in accordance with the method of concatenated, multi-parameter, mapped, fixed-point coding [5]. Here, a chromosome is a sequence of *k*-parts, each of them with **n** (encoding precision) genes. In the case of tuning of three controller parameters – K_p, T_i and T_d, the chromosome is a sequence of three parts. In the case of tuning of all the defined parameters – K_p, T_i, T_d, b, c and N, the chromosome is a sequence of six parts. The ranges of PID values are rationally chosen and it is true that the limitation will influence the results of the GA search; it is intended to obtain more stable, efficient and accurate solutions.

For the problem addressed, the ranges of the tuning parameters are considered as follows:

$K_p \in [0, inf]$,
$T_i \in [0, inf]$,
$T_d \in [0, inf]$,
$b \in [0, 5]$,
$c \in [0, 5]$ and $N \in [5, 1000]$.

Following a random initial choice, entire generations of such strings are readily processed in accordance with the basic genetic operators of selection, crossover and mutation. In particular, the selection process ensures that the successive generations of PID controller parameters, produced by the GA exhibit progressively improving behaviour with respect to some fitness measure.

Objective function

To evaluate the significance of the tuning procedure and controller performance, the integrated square error (I_{ISE}) criteria is used:

$$I_{ISE} = \int_0^T e(t)^2 \, dt \,, \tag{30}$$

In this case, error e is the difference between the set point and the estimated substrate concentration ($\gamma_{S_{sp}} - \gamma_S$). Based on [19], the $\gamma_{S_{sp}}$ is set to 0.1 g·l^{-1}.

5 Results and discussion

A control system based on PID controller to control the substrate concentration at 0.1 g·l^{-1} is designed. Using the considered objective function (Eq. (30)) series of tuning tests are performed. In the tuning procedure, the non-linear model of the considered process (Eqs. (1) - (5)), extended with the model of glucose correction (Eq. (28)), is applied. To obtain more realistic tests of the control systems applicability and of the tuning procedure performance, measurement noise and process noises are introduced (see Section 2). As a result of the GA tuning, the optimal PID controller settings are obtained. The results are presented in Table 2.

Table 2. Optimal controller parameters

PID controller		Digital PID controller	
Parameter	**Value**	**Parameter**	**Value**
K_p	0.0729	b_{i1}	0.0030
T_i	0.1377	b_d	0.2899
T_d	0.0324	a_d	0.3119
b	0.8087	b	0.8087
c	1.0158	c	1.0158
N	12.7543	N	12.7543
I_{ISE}	16.8187		

In Fig. 3, resulting dynamics of glucose concentration is presented. As it can be seen, the controller sets the control variable to the desire set point sufficiently fast, and keeps it to the end of cultivation process. The maximum deviation from the set point is 0.03 g·l^{-1} at time 15.85 h. Before the 14 h of cultivation, the maximum deviation is 0.01 g·l^{-1}. After 14 h, the measured and estimated uncorrected glucose concentrations are significantly decreased. It this case the noise/signal ratio in the measured glucose is increased which mainly affects the feedforward term of control signal. As a result, the oscillation of the control variable is observed (Fig. 3). In this moment, significant increase of the

specific growth rate leads to rapid increase of the control signal (Fig. 4). Nevertheless, the designed controller successfully keeps the glucose concentration close to the desired set point. Moreover, the control signal does not reach the actuator limitations and exhibits no oscillations which are undesired for the feed rate pump.

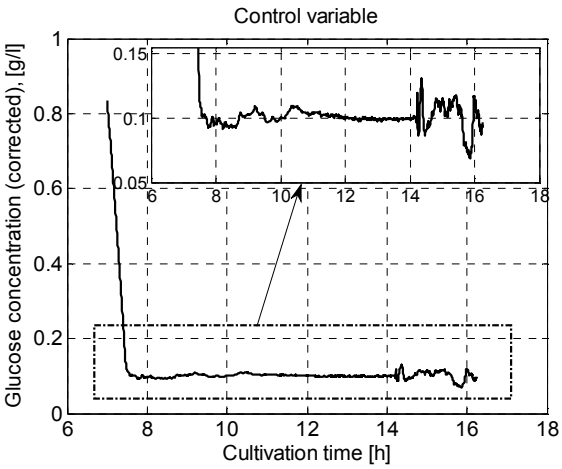

Fig. 3. Dynamics of corrected glucose concentration

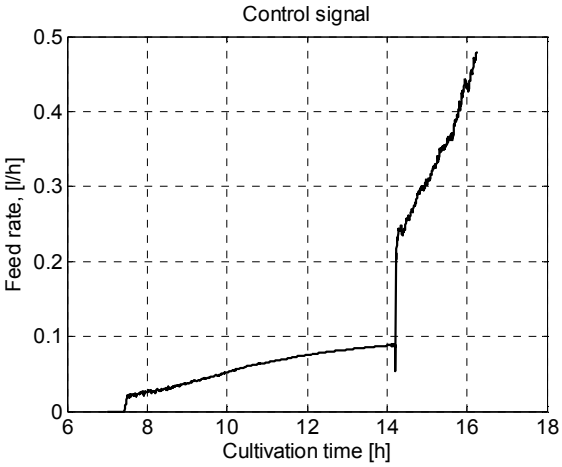

Fig. 4. Control signal

If the proposed correction is not used in the control algorithm, an increasing static error for the glucose concentration occurs. The desired glucose concentration (set point of 0.1 $g \cdot l^{-1}$) and the resulting one are shown in Fig. 5.

In Fig. 6, resulting dynamics of biomass concentration is presented. Good controller performance allows the cultivation process to continue for the maximum possible time (until the bioreactor volume reaches 2 l). As a result, at the end of the process high biomass concentration of 43.3 $g \cdot l^{-1}$ is obtained. The estimates of biomass and glucose concentration, as well as the estimates of maximum growth rate and bioreactor volume are presented respectively in Figs. 6 - 9.

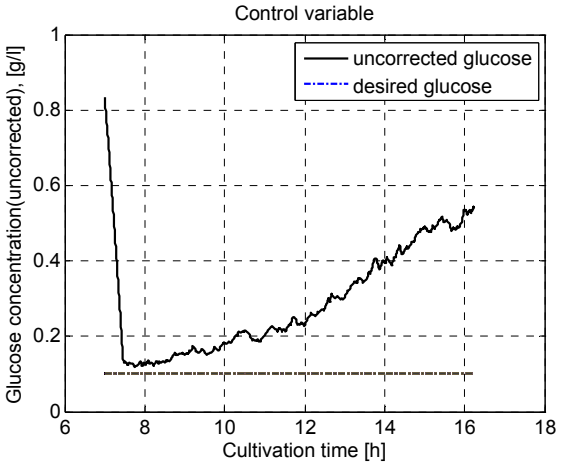

Fig. 5 Dynamics of uncorrected glucose concentration

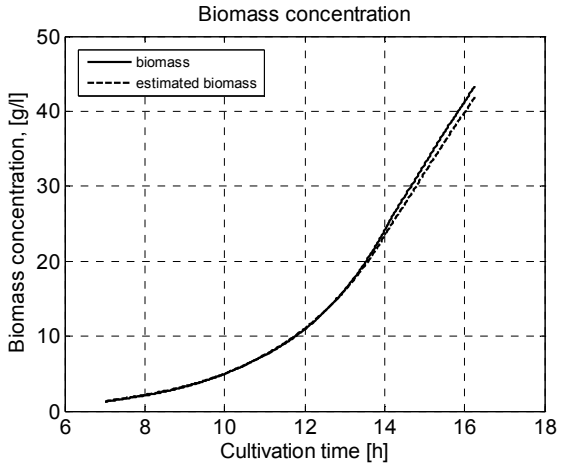

Fig. 6. Dynamics of biomass concentration and the estimated one

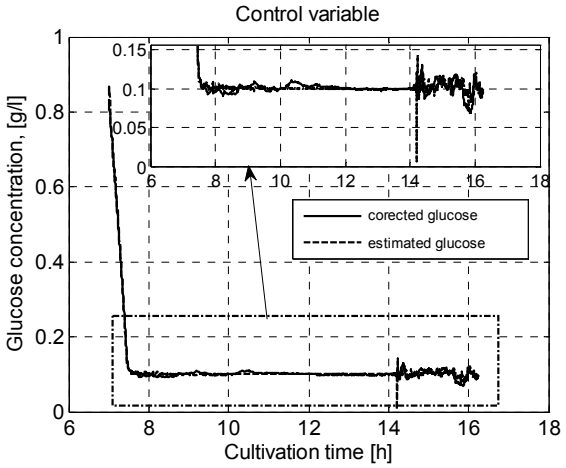

Fig. 7. Dynamics of corrected glucose concentration and the estimated one

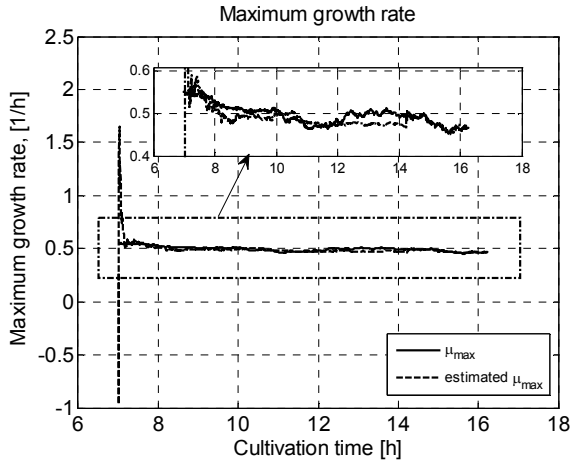

Fig. 8. Dynamics of maximum growth rate and the estimated one

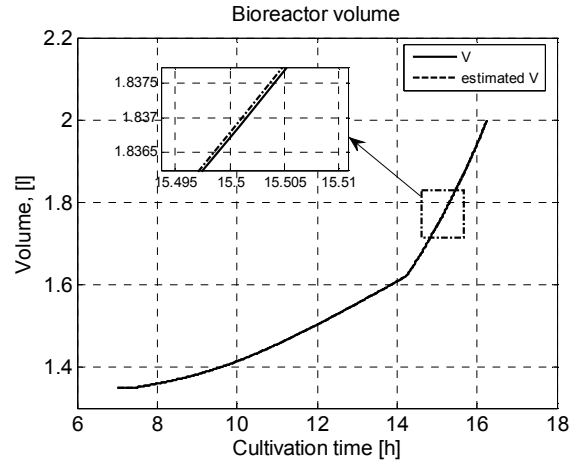

Fig. 9. Dynamics of bioreactor volume and the estimated one

The results show that the obtained estimates are close to the corresponding real signals. For example, the resulting error of glucose concentration estimation is shown in Fig. 10. Some statistical results for estimation errors are summarized in Table 3.

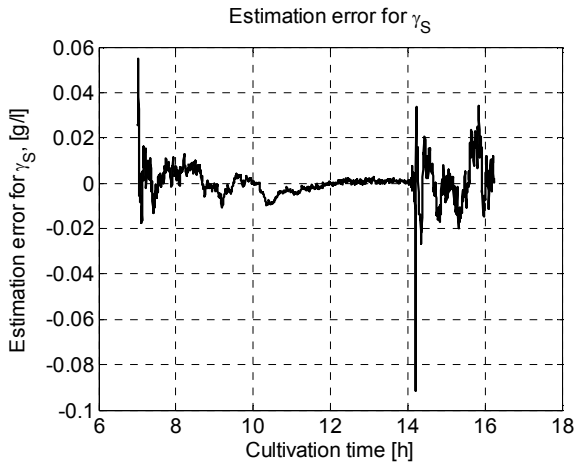

Fig. 10. Estimation error of glucose concentration

Table 3. Statistical results for estimation errors

Estimation error	$M\{\bullet\}$	$\sigma\{\bullet\}$
$\gamma_{S_{cor}} - \hat{\gamma}_{S_{cor}}$	3.126e-004	0.0073
$\gamma_X - \hat{\gamma}_X$	-0.221	0.4710
$\mu_{\max} - \hat{\mu}_{\max}$	-0.0043	0.0729
$V - \hat{V}$	2.168e-005	3.2155e-005

The values for the standard deviations of estimation errors are sufficiently small. The corresponding mean values are close to zero, which is indication for unbiased estimates. The maximum difference between glucose concentration and reference signal achieved here is 0.032 g·l^{-1}. The deviation from the set point is very small for the whole time period. In parallel, the maximum difference reported in [19] is 0.06 g·l^{-1} and it occurs in the second half of the process. The resulting standard deviation (σ_{γ_S}) and the mean value (M_{γ_S}) concerning control variable are:

σ_{γ_S} = 0.0071 g·l^{-1} and M_{γ_S} = 0.1006 g·l^{-1}

in this report and

σ_{γ_S} = 0.1513 g·l^{-1} and M_{γ_S} = 0.1306 g·l^{-1}

in [19].

6 Conclusion

In this article, the results of a designed feedforward feedback control system, based on the universal digital PID controller and extended Kalman filter are presented. The controller is used to control feed rate and to maintain glucose concentration at the desired set point for an *E. coli* MC4110 fed-batch cultivation process. A correction in the measured glucose concentration is proposed in order to affect the time delay of glucose measurement system. The correction is used to correct the estimates of glucose concentration obtained with the extended Kalman filter. For PID controller parameters GA are applied. As a result, an optimal PID controller settings are obtained and good closed-loop system performance is achieved. For a short time, the GA-tuned PID controller sets the glucose concentration and maintains it at the desired set point during the cultivation process. Sufficiently small values of the standard deviations of estimation errors are obtained. The estimation mean errors values are close to zero which is an indication for unbiased estimates. Furthermore, the designed control system allows carrying out the *E. coli* fed-batch cultivation process for the maximum permissible time in with respect to the maximum bioreactor volume. As a

result, the high biomass concentration of 43.3 g·l^{-1} is obtained at the end of the cultivation process. Finally, it is demonstrated that the genetic algorithms provide a simple, efficient and accurate approach to PID controllers tuning. Moreover, obtained results show that genetic algorithm tuning can be considered as an effective methodology for achieving high quality and better performance of the designed control system.

Nomenclature

γ_X	the concentration of biomass, g·l^{-1}
γ_S	concentration of substrate (glucose), g·l^{-1}
Q	feed rate, l·h^{-1}
V	bioreactor volume, l
γ_{in}	substrate concentration of the feeding solution, g·l^{-1}
$\mu_{\max}$	maximum growth rate, h^{-1}
k_S	saturation constant, g·l^{-1}
$Y_{S/X}$	yield coefficient, -
η_{γ_X}	biomass concentration process noises, g·l^{-1}
η_{γ_S}	substrate concentration process noises, g·l^{-1}
η_Q	feed rate process noises, l·h^{-1}
$\eta_{\mu_{\max}}$	maximum growth rate process noises, h^{-1}
$\xi(t)$	measurement noise, g·l^{-1}
$u_{fb}(s)$	control variable, l·h^{-1}
$r(s)$	reference signal, g·l^{-1}
$\gamma_S(s)$	process output, g·l^{-1}
K_p	proportional gain, -
T_i	integral time, h
T_d	derivative time, h
b, c	set point weight coefficients, -
T_d / N	low-pass first order filter of D-term time-constant, h
T_0	sample time, h
k	discrete time, -
$\hat{\mathbf{X}}(\cdot)$	the estimate of $\mathbf{X}(\cdot)$, g·l^{-1}
$\hat{\gamma}_S(\cdot)$	the estimate of glucose concentration $\gamma_S(\cdot)$, g·l^{-1}
$\mathbf{K}_{\text{EKF}}(\cdot)$	the EKF gain, -
$\mathbf{P}(\cdot)$	the covariance matrix
$\mathbf{F}(k)$	the Jacobian of nonlinear function $f_d(\mathbf{X}(k))\big\|_{\mathbf{X}(k)=\hat{\mathbf{X}}(k)}$
$\mathbf{D}_{\boldsymbol{\eta}\mathbf{d}}$	the covariance matrix of discrete-time process noise $\boldsymbol{\eta}_d(k) = T_0\boldsymbol{\eta}(t)$
D_ξ	the covariance of measurement noise
e	error, g·l^{-1}
t	time, h
T	end time of the cultivation, h
$\gamma_{S_{sp}}$	set point of glucose concentration, g·l^{-1}
$\sigma\{\bullet\}$	standard deviation
$M\{\bullet\}$	mean value
$D\{\bullet\}$	variance

Acknowledgements:
The investigations are partially supported by the Bulgarian National Science Fund, Grants DID 02/29 "Modeling Processes with Fixed Development Rules (ModProFix)" and DMU 02/4 "High Quality Control of Biotechnological Processes with Application of Modified Conventional and Metaheuristics Methods".

References:

[1] Lange, U., Vassileva, S., Zipser, L., Controllable Shock and Vibration Dampers Based on Magnetorheological Fluids, In Electrical and Computer Engineering Series – A Series of Reference Books and Textbooks "Advances in Systems Theory, Mathematical Methods and Applications", Ed. Alexander Zemliak and Nikos E. Mastorakis, WSEAS Press, 2002, pp. 245-250.

[2] Ventzas, D., Theodorellos, G., Microprocessor-based Simulation of the Intelligent Integrator, International Workshop on Microprocessor Control of Small Motors, IEEE-IES & AEI, Padova, Italy, July 14-15 1986.

[3] Kumar, S. M., Giriraj, R., Jain, N., Anantharaman, V., Dharmalingam, K. M. M., Sheriffa, B., Genetic Algorithm based PID Controller Tuning for a Model Bioreactor, *Indian Chemical Engineer*, Vol. 50(3), 2008, pp. 214-226.

[4] Holland, J. H., *Adaptation in Natural and Artificial Systems*, 2nd Ed. Cambridge, MIT Press, 1992.

[5] Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley Longman, London, 2006.

[6] Parker, B. S., *Demonstration of Using Genetic Algorithm Learning*, Information Systems Teaching Laboratory, 1992.

[7] Oonsivilai, A., Oonsivilai, R., A Genetic Algorithm Application in Natural Cheese Products, *WSEAS Transactions on Systems*, Vol. 8(1), 2009, pp. 44-54.

[8] Chang, W.-Y., Application of Grey Clustering Approach and Genetic Algorithm to Partial Discharge Pattern Recognition, *WSEAS Transactions on Systems*, Vol. 8(12), 2009, pp. 1273-1283.

[9] Neubauer, A., Theory of the Simple Genetic Algorithm with α-Selection, Uniform Crossover and Bitwise Mutation, *WSEAS Transactions on Systems*, Vol. 9(9), 2010, pp. 989-998.

[10] Neri, F., Learning and Predicting Financial Time Series by Combining Evolutionary Computation and Agent Simulation, *Transactions on Computational Collective Intelligence*, Vol. 6, Springer, Heidelberg, 2011, pp. 202-221.

[11] Kim, D. H., Comparison of PID Controller Tuning of Power Plant Using Immune and Genetic Algorithm. Measurements and applications, *Proceedings of Computational Intelligence for Measurement Systems and Applications*, Ligano, Switzerland, 2003, pp. 169-174.

[12] Kim, D. H., Park, J. I., Intelligent PID Controller Tuning of AVR System Using GA and PSO, *Lecture Notes in Computer Science*, Vol. 3645, 2005, pp. 366-375.

[13] Kim, J. S., Kim, J.-H., Park, J.-M., Park, S.-M., Choe, W.-Y., Heo, H., Auto Tuning PID Controller Based on Improved Genetic Algorithm for Reverse Osmosis Plant, *World Academy of Science, Engineering and Technology*, Vol. 47, 2008, pp. 384-389.

[14] Gundogdu, O., Optimal-tuning of PID Controller Gains Using Genetic Algorithms, *Journal of Engineering Sciences*, Vol. 11(1), 2005, pp. 131-135.

[15] Mahony, T. O., Downing, C. J., Fatla, K., Genetic Algorithm for PID Parameter Optimization: Minimizing Error Criteria, *Proceedings of Process Control and Instrumentation*, University of Stracthclyde, 2000, pp. 148-153.

[16] Wang, Q., Spronck, P., Tracht, R., An Overview of Genetic Algorithms Applied to Control Engineering Problems, *Machine Learning and Cybernetics*, Vol. 3, 2003, pp. 1651-1656.

[17] Krohling, R. A., Rey, J. P., Design of Optimal Disturbance Rejection PID Controllers Using Genetic Algorithm, *IEEE Trans. Evol. Comput*., Vol. 5, 2001, pp. 78-82.

[18] Roeva, O., Pencheva, T., Hitzmann, B., Tzonkov, St., A Genetic Algorithms Based Approach for Identification of *Escherichia coli* Fed-batch Fermentation, *Int. J. Bioautomation*, Vol. 1, 2004, pp. 30-41.

[19] Arndt, M., Hitzmann, B., Feed Forward/feedback Control of Glucose Concentration during Cultivation of *Escherichia coli*, *Proceedings of the 8th IFAC Int Conf on Comp Appl in Biotechn*, 2001, pp. 425-429.

[20] Astrom, K., Hagglund, T., *PID Controllers* (2nd ed.), Instrument Society of America, Research Triangle Park, NC, USA, 1995.

[21] Astrom, K., Hagglund, T., *Advanced PID Control*, ISA – The Instrumentation, System and Automation Society, 2006.

[22] Garipov, E., PID Controllers, *Automatics and Informatics*, Vol. 3, 2006. (in Bulgarian)

[23] Kotsiantis, S., Kanellopoulos, D., Discretization Techniques: A Recent Survey, *GESTS International Transactions on Computer Science and Engineering*, Vol. 32(1), 2006, pp. 47-58.

[24] Heath, M. T., *Scientific Computing. An Introductory Survey*, (2nd ed.), McGraw-Hill, New York, 2002.

[25] Johnson, M., Moradi, M. H., *PID Control: New Identification and Design Methods*, Springer, 2005.

[26] O'Dwyer, A., *Handbook of PI and PID Controller Tuning Rules*, Imperial College Press, 2006.

[27] Roeva, O., Improvement of Genetic Algorithm Performance for Identification of Cultivation Process Models, *Advanced Topics on Evolutionary Computing, Book Series: Artificial Intelligence Series – WSEAS*, 2008, pp. 34-39.

[28] Roeva, O., Trenkova, T., Cultivation Process Optimization based on Genetic Algorithms, *Journal of International Scientific Publication: Materials, Methods & Technologies*, Vol. 4(2), 2010, pp. 121-136.