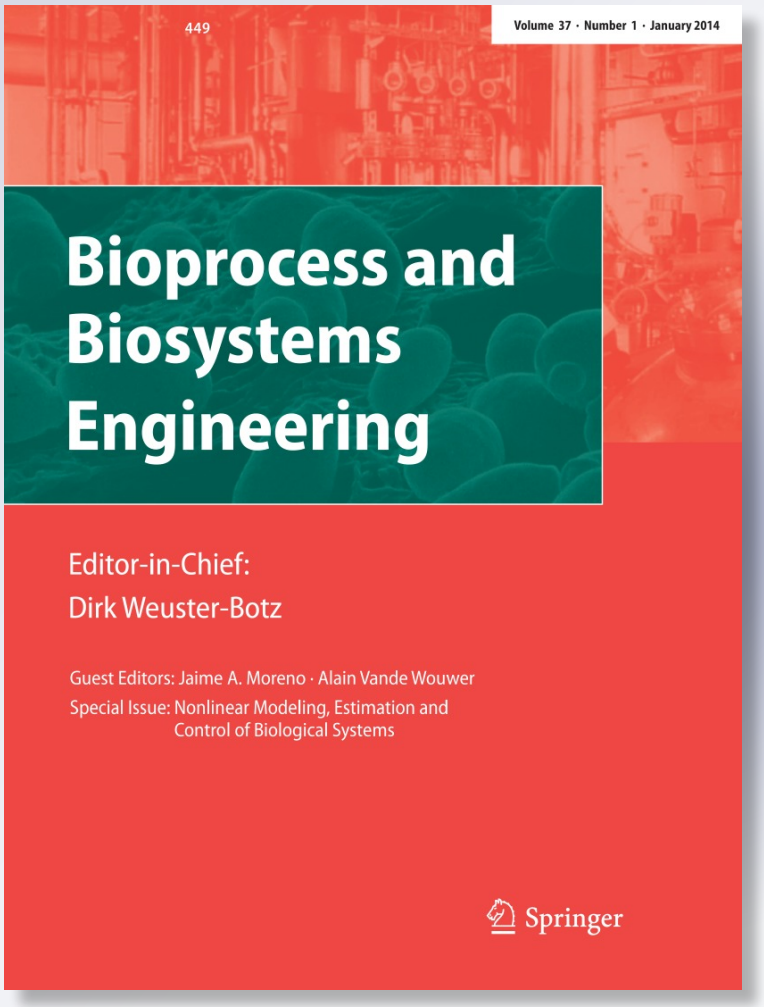
Numerical aspects of optimal control of penicillin production

Matej Pčolka & Sergej Čelikovský

Bioprocess and Biosystems Engineering

ISSN 1615-7591
Volume 37
Number 1

Bioprocess Biosyst Eng (2014) 37:71-81
DOI 10.1007/s00449-013-0929-z



Springer

Bioprocess Biosyst Eng (2014) 37:71–81
DOI 10.1007/s00449-013-0929-z

ORIGINAL PAPER

Numerical aspects of optimal control of penicillin production

Matej Pčolka · Sergej Čelikovský

Received: 29 September 2012 / Accepted: 12 February 2013 / Published online: 20 March 2013

Abstract Since their discovery, fermentation processes have gone along not only with the industrial beverages production and breweries, but since the times of Alexander Fleming, they have become a crucial part of the health care due to antibiotics production. However, complicated dynamics and strong nonlinearities cause that the production with the use of linear control methods achieves only suboptimal yields. From the variety of nonlinear approaches, gradient method has proved the ability to handle these issues—nevertheless, its potential in the field of fermentation processes has not been revealed completely. This paper describes constant vaporization control strategy based on a double-input optimization approach with a successful reduction to a single-input optimization task. To accomplish this, model structure used in the previous work is modified so that it corresponds with the new optimization strategy. Furthermore, choice of search step is explored and various alternatives are evaluated and compared.

Keywords Optimal control · Nonlinear systems · Fermentation process · Gradient method optimization · Antibiotics production

M. Pčolka (✉)
Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Karlovo náměstí 13, 121 35 Prage 2, Czech Republic
e-mail: pcolkmat@fel.cvut.cz

S. Čelikovský
Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Prague 8, Czech Republic
e-mail: celikovs@utia.cas.cz

Introduction

Rapid increase of the industrial productivity of antibiotics that might be witnessed during the last few decades is basically owed to a massive improvement of production technologies rather than to a sophisticated control background. As a consequence, only suboptimal operation manners have been involved with final product concentrations deep below maximum reachable values. The important next step is to consider and carefully analyze all advantages and disadvantages of the used control strategy. A wide variety of ways how to operate the input feed flow (which influences the formation of the final product especially by the amount of the substrate nutrient supplied to system through it) has been discussed in literature so far. As an initial attempt, one can consider indirect feedback methods for nutrient feeding based on pH or dissolved oxygen measurements [1]—the substrate concentration is then maintained at predetermined setpoint by either a simple open-loop controller [2] or an on/off [3] or a PID type controller, followed by fuzzy approaches which appeared in the 1990s [4] and have been revitalized at the beginning of the millennium [5]. However, the most impressive results have been reached using model predictive control (MPC) approach. Several studies describing the MPC control of bioprocess, in general [6–10], and penicillin production, in particular [11, 12], can be found. The main drawback of this method is the fact that it is usually performed either with an approximately or exactly linearized mathematical model of the controlled process. Approximate linearization performed at certain operating point [7] can be invalid for operating points far away from the original one (and it is known that the operating points range varies a lot during the cultivation). Moreover, no stability assumptions can be made for closed-loop control

based on the approximate models obtained at each step and even one unstable model obtained by approximate linearization can degrade the MPC performance vastly. Exact linearization blows all these problems away—unfortunately, in the area of fermentation processes, the existence of exact linearization is rather rare and occasional. Therefore, a proper alternative is needed—gradient descent method which has already proved encouraging results in various research areas [13–15] is a strong candidate as it can handle even a nonlinear process model very effectively. The crucial points for this model based method are the availability of a mathematical model describing the biochemical process and determination of an adequate cost functional to be optimized.

The aim of this paper is to continue the previous efforts of the authors in [16], where they consider the single-input model with the input being the feed flow. The first attempts to consider a non-nutrient input and use different strategies for the double-input model were presented in [17], while the comparison of both the so-called quasi-double-input and the true-double-input strategies was performed in [18]. Based on that, the present paper goes deeper into the strategy used in [17] and studies further its efficiency both from the biotechnological and the numerical point of view. In the current paper, gradient search step choice is discussed and three alternative families are provided: (i) fixed step family, (ii) parabola-minimizing step family, and (iii) general-curve-minimizing step family. Each of the mentioned families contains more members whose results are later compared with respect to the following criteria: optimality, iterations-to-converge and time-to-converge.

The paper is organized as follows: "Model of the fermentation process" introduces nonlinear dynamical model of the fermentation process which is used for the optimization purposes. The penicillin cultivation is chosen to represent the fermentation processes, modification of the previously used model (which is crucial for the use of new control strategy introduced later in this paper) is explained. In "Optimal control design", the optimization issues of the final product concentration maximization including the constraints specification are formulated. The gradient method is introduced, its theoretical background is clarified and having done the necessary problem order reduction, the used control strategy is proposed. In "Optimization results", results of constant vaporization strategy are presented, compared to those obtained using strategy presented earlier and discussed. "Choice of gradient search step" introduces search step families and brings a brief description of particular family members, while "Results with enhanced search step choice" summarizes the results of the numerical experiments for different step choices and comments upon them. "Conclusions" concludes the paper.

Model of the fermentation process

Let us consider a penicillin cultivation [12, 16, 19] described by the following model:

$$\begin{aligned}
\frac{\mathrm{d}V}{\mathrm{d}t} &= u - V\lambda\left(\mathrm{e}^{w\frac{T_{\mathrm{opt}}-T_{\mathrm{f}}}{T_{\mathrm{b}}-T_{\mathrm{f}}}}-1\right),\\
\frac{\mathrm{d}C_{\mathrm{X}}}{\mathrm{d}t} &= (\mu - K_{\mathrm{D}})C_{\mathrm{X}} - \frac{\mathrm{d}V}{\mathrm{d}t}\frac{C_{\mathrm{X}}}{V},\\
\frac{\mathrm{d}C_{\mathrm{S}}}{\mathrm{d}t} &= -\sigma C_{\mathrm{X}} + \frac{C_{\mathrm{S,in}}u}{V} - \frac{\mathrm{d}V}{\mathrm{d}t}\frac{C_{\mathrm{S}}}{V},\\
\frac{\mathrm{d}C_{\mathrm{P}}}{\mathrm{d}t} &= \pi C_{\mathrm{X}} - K_{\mathrm{H}}C_{\mathrm{P}} - \frac{\mathrm{d}V}{\mathrm{d}t}\frac{C_{\mathrm{P}}}{V}.
\end{aligned} \tag{1}$$

Here, V (l) refers to cultivation broth volume, C_{X} (gl^{-1}) represents biomass concentration, C_{S} (gl^{-1}) stands for the limiting substrate concentration (let us consider carbon to be the limiting substrate), and C_{P} (gl^{-1}) represents the final product (penicillin) concentration. The substrate feed flow rate u (lh^{-1}) is the operated input.

Parameters λ (h^{-1}), and w ($-$) are specific vaporization constants, T_{opt} (K) represents empirically obtained optimal operational temperature (see [20]), and T_{f} (K) and T_{b} (K) refer to the freezing and the boiling temperature of the broth, respectively, which are considered to be the same as those of the water [12].

A simple constant term K_{D} (h^{-1}) models biomass death kinetics, while the total of specific biomass growth rate μ (h^{-1}) and the specific production rate π (h^{-1}) weighted by biomass-on-substrate yield coefficient $Y_{\mathrm{X/S}}$ and the product-on-substrate yield coefficient $Y_{\mathrm{P/S}}$ gives specific substrate consumption rate σ (h^{-1}):

$$\sigma = Y_{\mathrm{X/S}}^{-1}\mu + Y_{\mathrm{P/S}}^{-1}\pi.$$

In this paper, Contois kinetics of the biomass growth [21] and Haldane kinetics [22] of the product formation are considered, which results in the following expressions for the μ and π:

$$\mu = \mu_{\max}\frac{C_{\mathrm{S}}}{K_{\mathrm{X}}C_{\mathrm{X}} + C_{\mathrm{S}}},\ \pi = \pi_{\max}\frac{C_{\mathrm{S}}}{K_{\mathrm{P}} + C_{\mathrm{S}} + C_{\mathrm{S}}^2/K_{\mathrm{I}}}, \tag{2}$$

where $\mu_{\max}$ (h^{-1}), $\pi_{\max}$ (h^{-1}) are the maximum specific growth and production rates, K_{X} ($-$) is the Contois saturation constant, K_{P} (gl^{-1}) is product formation saturation constant and K_{I} (gl^{-1}) is inhibition constant for product formation.

Input substrate concentration $C_{\mathrm{S,in}}$ (gl^{-1}) reflects the effect of the input flow u on the substrate concentration C_{S}. Finally, penicillin hydrolysis is modeled by a degradation

constant K_H (h^{-1}). At this point, more interested readers are referred to [19] and [12] where the model is described in more detail.

Now, let us assume that the cultivation volume V can not only be increased by exogenous input u, but a volume withdrawal can be performed as well. This requires a new input variable to be introduced and the volume differential equation changes into:

$$\frac{dV}{dt} = u_1 - V\lambda\left(e^{w\frac{T_{opt}-T_f}{T_b-T_f}} - 1\right) - u_2, \tag{3}$$

where u_1 corresponds with the old input variable, and u_2 (lh^{-1}) stands for volume withdrawal. The resulting scheme of the cultivation tank is shown in Fig. 1.

From the practical point of view, the introduction of the volume withdrawal brings several advantages—firstly, the engineer can control the tank volume and the tank overflow can be prevented. In industrial practice, this is often the main reason of introducing the effluent flow; however, the withdrawal is usually controlled ad hoc (if the volume reaches chosen level, certain part of the broth is withdrawn). The approach presented in this paper and described in detail in the following section takes the volume withdrawal directly into account and exploits it in favor of final product concentration maximization. Secondly, all the state variables are present in the withdrawn broth which can be exploited for state variables measurements (which in many cases of cultivations is performed manually) and cultivation analysis.

This little change of volume differential, however, effects the differential equations of other state variables as well. Let us remind that the state variables are in form of concentration which cannot be increased nor decreased by volume withdrawal. Therefore, terms including volume differential should be modified as follows: $dV/dt \rightarrow dV/dt + u_2$.

For the needs of optimization and to follow the conventional notation, let us rewrite the extended model (1) into the ordinary form using $x^T = [x_1, \ldots, x_4] = [V, C_X, C_S, C_P]$ and $u^T = [u_1, u_2]$:

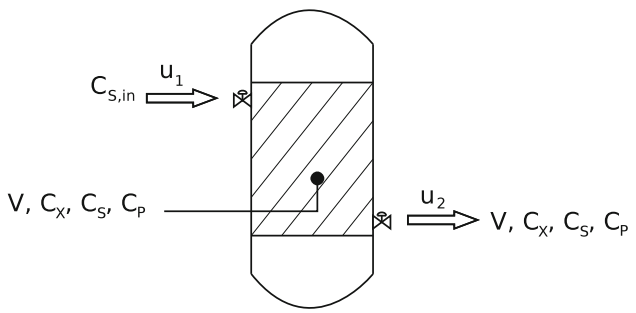


Fig. 1 Scheme of the cultivation tank

$$\begin{aligned}
\dot{x}_1 &= u_1 - \lambda(e^{w\frac{T_{opt}-T_f}{T_b-T_f}} - 1)x_1 - u_2,\\
\dot{x}_2 &= \left(\mu_{max}\frac{x_3}{K_X x_2 + x_3} - K_D\right)x_2\\
&\quad - \left(u_1 - \lambda\left(e^{w\frac{T_{opt}-T_f}{T_b-T_f}} - 1\right)x_1\right)\frac{x_2}{x_1},\\
\dot{x}_3 &= -\left(\frac{\mu_{max}}{Y_{X/S}}\frac{x_3}{K_X x_2 + x_3} + \frac{\pi_{max}}{Y_{P/S}}\frac{x_3}{K_P + x_3 + x_3^2/K_I}\right)x_2\\
&\quad + \frac{C_{S,in}u_1}{x_1} - \left(u_1 - \lambda\left(e^{w\frac{T_{opt}-T_f}{T_b-T_f}} - 1\right)x_1\right)\frac{x_3}{x_1},\\
\dot{x}_4 &= \pi_{max}\frac{x_3}{K_P + x_3 + x_3^2/K_I}x_2 - K_H x_4\\
&\quad - \left(u_1 - \lambda\left(e^{w\frac{T_{opt}-T_f}{T_b-T_f}} - 1\right)x_1\right)\frac{x_4}{x_1}.
\end{aligned} \tag{4}$$

For the better comprehension of the model, let us introduce a brief description of the model and an explanation of the phenomena typical for the penicillin cultivation.

The first differential equation describes the change of the volume profile. The increase of the volume happens due to the presence of the first input u_1 and it is decreased either applying the second input u_2 or in a very natural way due to the vaporization expressed by the vaporization term.

Like every other living organism, the biomass reproduces and dies—in the model (4), the increase of the biomass concentration is described by the reproduction term represented by biomass growth rate μ, while the decrease of biomass concentration follows from biomass death modeled by a constant death rate K_D. Except of these physiological ways of growth and decay, the biomass concentration is affected in a slightly "artificial" way—thanks to the vaporization, the biomass concentration increases while the volume increase (caused by the feed poured into the tank) leads to its decrease.

The "fuel" (essential nutrient) which is consumed by the "driving engine" of the whole bioprocess (biomass) is represented by the substrate concentration. It is crucial not only for keeping biomass alive but also for the product formation. Both of these consumption phenomena are described by the third state differential equation including the yield constants $Y_{X/S}$ and $Y_{P/S}$. Moreover, dual impact of input feed flow on substrate concentration can be observed—the input feed flow increases the substrate concentration (via qualitative constant of the feed $C_{S,in}$) and, on the other hand, the level of the substrate concentration in the broth is decreased due to dilution. Following from the concentration character of the third state variable, the vaporization term is included in its differential equation as well.

Product concentration (the most attractive variable from the industrial point of view) is increased at production rate π which, however, is completely different from the growth rate μ. This reflects different phases of the microorganisms

life—at certain phase, either the biomass reproduction or the product formation is preferred. Hydrolysis of the penicillin is modeled by constant term K_H, while the last two terms are related to the concentration nature of the fourth state variable.

The model (4) is adapted to the newly proposed constant vaporization strategy described in the next section, and it is further used in the optimization procedure as the representant of the controlled system behavior.

Optimal control design

In [16], an optimal feeding strategy coming out of a projected gradient method has been introduced. Theoretical complication given by the state dependence of the input saturation has been successfully addressed, and assumption on sufficiently large cultivation tank volume has been made. However, in industrial application, the cultivation tank may be filled up with a such large initial volume that application of the computed input feed flow rate leads to the tank overflow in short horizon. A perspective-offering solution to this problem has been tackled in the previous section. Here, we propose a way to operate the second input which can bring interesting results improvement.

Constant vaporization strategy

Applying another exogenous input u_2, one can avoid tank overflow, yet another problem occurs. A thoughtful reader has surely already noticed that having introduced two input variables, the first differential equation of mathematical description of the system does not comply with physical laws. It can be shown that at certain point the volume can reach zero value and further withdrawal can theoretically cause negative volume, which is physically impossible. One way of avoiding this is to set a dynamical constraint on the second input u_2 which ensures that at the point of zero volume V, the withdrawal does not exceed the inlet flow. However, looking at the issue from the engineering point of view, it is not either convenient to decrease the volume below certain too low value, as the final product amount equals to concentration C_P multiplied by the volume V.

Let us introduce an idea leading to a strategy solving the sketched negative volume difficulty. It consists in an assumption that the second input u_2 is used to compensate the effect of the first input u_1 on the volume V. From the first differential equation of model (4), it is obvious that the volume is affected by feed flow rate u_1, volume withdrawal rate u_2 and by natural vaporization described by the middle term. The key idea of the constant vaporization strategy is that we require the volume to be just naturally vaporizing without any other dynamical response to the exogenous signals—from the first differential equation of the model (4) which (in agreement with the requirement for constant vaporization of the broth) should be equal to vaporization term only, the second input u_2 can be calculated directly as $u_2 = u_1$, which results in $\dot{x}_1 = -K_{vap}x_1$ where K_{vap} is the overall vaporization constant, $K_{vap} = \lambda(\exp(w\frac{T_{opt}-T_f}{T_b-T_f}) - 1)$. Let us note that the sketched idea of double-input problem simplification is also described in [18] where except of quasi-double-input strategies, the true-double-input strategies performing complete optimization with both inlet and outlet flow inputs are described and compared.

Following the above mentioned strategy, the inlet flow stays the only optimization variable and the system description of the original process given by (4) changes into the following model with reduced input set:

$$\begin{aligned}
\dot{\xi}_1 &= -K_{vap}\xi_1,\\
\dot{\xi}_2 &= \left(\mu_{max}\frac{\xi_3}{K_X\xi_2+\xi_3} - K_D\right)\xi_2 - \left(\upsilon - K_{vap}\xi_1\right)\frac{\xi_2}{\xi_1},\\
\dot{\xi}_3 &= -\left(\frac{\mu_{max}}{Y_{X/S}}\frac{\xi_3}{K_X\xi_2+\xi_3} + \frac{\pi_{max}}{Y_{P/S}}\frac{\xi_3}{K_P+\xi_3+\xi_3^2/K_I}\right)\xi_2\\
&\quad + \frac{C_{S,in}\upsilon}{\xi_1} - \left(\upsilon - K_{vap}\xi_1\right)\frac{\xi_3}{\xi_1},\\
\dot{\xi}_4 &= \pi_{max}\frac{\xi_3}{K_P+\xi_3+\xi_3^2/K_I}\xi_2 - K_H\xi_4 - \left(\upsilon - K_{vap}\xi_1\right)\frac{\xi_4}{\xi_1}.
\end{aligned} \tag{5}$$

State vector ξ corresponds to the original state vector, $\xi = [\xi_1, \ldots, \xi_4]^T = [V, C_X, C_S, C_P]^T$ while υ represents inlet feed flow rate.

Model (5) supplemented by the corresponding model parameters (Table 1) provides an engineer with a tool to

Table 1 Model parameters

Parameter	Value
μ_{max}	0.11
π_{max}	0.004
K_P	0.1
$Y_{X/S}$	0.47
K_D	0.0136
K_X	0.06
K_H	0.01
$Y_{P/S}$	1.2
$C_{S,in}$	500
K_{vap}	6.23×10^{-4}
K_I	0.1
T_{opt}	298
T_f	273
T_b	373

design the optimal control minimizing a properly chosen criterion.

Optimization task formulation

From the optimization point of view, penicillin production optimization can be viewed as fixed initial state, free time interval and free final state issue. Without any loss of generality and due to upper cultivation duration constraint, let us now consider multiple optimization routines with fixed time intervals of length $t_{\text{end}} \in \{200, 300, 400, 500\}$. This helps us to simplify the optimization procedure and avoid difficulties with general time interval solution.

For the optimization purpose, the objective functional reflecting the optimization effort needs to be formulated mathematically. From the industrial point of view, two quite antagonistic goals can be chosen—both the quantity (represented by the amount of the product) and the quality of the final product (represented by its concentration in the cultivation broth) can be desired to be maximized. It can be intuitively seen that following only one of these optimization directions, two extremes are reached neither of which is preferable. Maximization of product amount (without any concentration check) can end up with an extremely large volume containing only a very low level of penicillin concentration, while maximization of concentration (without any volume limitation) can lead to a highly concentrated tiny-volumed broth. Without doubt, the quality of the final product is the factor affecting the duration of the product post-processing and subsequently also the efficiency of the whole industrial process crucially—the more concentrated the broth is, the shorter post-processing procedure is needed and, therefore (assuming a very common situation in the industrial practice with multiple cultivation tanks, but only a limited number of post-processing machines available), the cultivation can be repeated more frequently which can positively influence the overall productivity. Moreover, the maximization of the product concentration is very often directly requested in the industrial practice—if some product amount is guaranteed, the industrial companies are usually interested in obtaining the product of highest possible quality. Taking the high-quality-product requirements into account, concentration maximization (also considered in [12, 23–25] and many other papers) is chosen to be the preferred optimization criterion. However, it has been already mentioned that the absence of volume limitation can results in a small volume—these issues are in detail handled in [18] where in the true-double-input cases, the volume constraints are applied. In the case considered in this paper, the effect of vaporization phenomena is not critical enough to degrade the control performance and therefore, no volume limitations need to be considered.

As the main goal is to maximize the final product concentration, the following criterion in the Mayer form is formulated:

$$\mathcal{J} = -\xi_4(t_{\text{end}}), \tag{6}$$

where $\mathcal{J}$ denotes the criterion for the constant vaporization strategy.

Regarding state optimization constraints, it can be shown that the model (5) satisfies physical constraints (state variables nonnegativity) and no further attention is necessary to be paid to low state constraints. Moreover, constant vaporization strategy eliminates the need for upper volume constraint handling. Thus, input saturation constraint $0 \le \upsilon \le U_{\max}$ and input piecewise constant character $\mathrm{d}\upsilon/\mathrm{d}t = 0$ for $ml \le t < (m+1)l$, $m = 0, 1, \ldots$ (accomplished by sampling of the inputs υ with sampling period $l = 4$ h) are the only static constraints related to this optimization task.

Having properly defined the system equations, the input constraints and the objective functional, the optimization problem for $t \in [t_0, t_{\text{end}}]$ (without any loss of generality, let us consider $t_0 = 0$ h) can be summarized:

$$\upsilon^*(t) = \arg\min_{\upsilon(t)} \mathcal{J}(\xi(t)) \tag{7}$$

such that the following constraints hold:

$$\begin{aligned} \dot{\xi}(t) &= f(\xi(t), \upsilon(t)), \\ \xi(t_0) &= \xi_0, \\ 0 &\le \upsilon(t) \le U_{\max}. \end{aligned} \tag{8}$$

Here, $f(\xi(t), \upsilon(t))$ refers to the model (5). The values of ξ_0 and $U_{\max}$ are summarized in Table 2.

Nonlinear gradient method

This method belongs to the family of the optimal control methods [26]. For the problem stated by (7) and the constraints given in the form of (8), the optimal input υ^* is searched iteratively. First of all, the initial input vector υ_0 is estimated (in our case, zero vectors have been chosen). Then, the following procedure is applied:

$$\upsilon^*_{k+1} = \upsilon^*_k - \alpha \frac{\partial \mathcal{J}}{\partial \upsilon}, \tag{9}$$

Table 2 Optimization constraints

Parameter	Value
$\xi_{1,0}$	To be specified
$\xi_{2,0}$	1.5
$\xi_{3,0}$	6
$\xi_{4,0}$	0
$U_{\max}$	0.05

where $k = 0, 1, 2, \ldots$ is the number of the iteration, and α is the search step parameter whose choice is described later. Here it should be noted that direct calculation of $\partial\mathcal{J}/\partial\upsilon$ is quite complicated due to the fact that ξ_4 depends on υ via a differential equation. Therefore, let us rather introduce Hamiltonian $\mathcal{H}$ in this general form:

$$\mathcal{H} = L + p^{\mathrm{T}} f. \tag{10}$$

Here, L represents the integral penalty of the optimized criterion, f refers to the model (5), and p is the adjoint state vector solved back in time. As our criterion $\mathcal{J}$ does not contain the integral penalty, the Hamiltonian turns into the following form:

$$\mathcal{H} = p^{\mathrm{T}} f. \tag{11}$$

To compute the gradient of (6) with respect to $\upsilon(t)$, set first:

$$\begin{aligned} -\dot{p} &= \frac{\partial \mathcal{H}}{\partial \xi}, \\ \dot{\xi} &= \frac{\partial \mathcal{H}}{\partial p}, \\ \xi(t_0) &= \xi_0, \\ p(t_{\mathrm{end}}) &= -\left(\frac{\mathrm{d}\phi}{\mathrm{d}\xi}|_{t=t_{\mathrm{end}}}\right), \end{aligned} \tag{12}$$

where ϕ is the terminal term of the optimization criterion. In our case, $\phi = -\xi_4(t_{\mathrm{end}})$ from which it follows $p(t_{\mathrm{end}}) = [0, 0, 0, 1]^{\mathrm{T}}$. It can be shown (mathematically rigorous proof is beyond the scope of this paper) that $\partial\mathcal{J}/\partial\upsilon = -\partial\mathcal{H}/\partial\upsilon$—thus, gradient $\partial\mathcal{H}/\partial\upsilon$ can be used in iterative procedure (9), which changes into:

$$\upsilon^*_{k+1} = \upsilon^*_k + \alpha\frac{\partial \mathcal{H}}{\partial \upsilon}. \tag{13}$$

At this moment, a constant search step parameter has been chosen $\alpha = 0.002$. Examination of another step choices is provided in the following section. Input saturation constraint is handled by mapping the iterated input vectors υ_{k+1} on an admissible input sets $\Upsilon_{\mathrm{admiss}} = \{\upsilon, 0 \le \upsilon \le U_{\max}\}$ by a simple saturation. Requirement of piecewise constant nature of the input υ is satisfied by sampling with sampling period $l = 4$ h.

The iterative procedure described by (13) terminates at the moment when the improvement obtained at the $(k + 1)$-st iteration is less than a chosen tolerance compared to the k-th optimization iteration result.

Optimization results

In this section, results obtained by the constant vaporization strategy are presented and compared to those obtained by the original one-input gradient method optimization (CG) presented in [16] and constant volume strategy which instead of natural vaporization keeps the volume constant. The latter one is described in [17] and [18]. The optimization results have been simulated with the penicillin cultivation model in MATLAB environment.

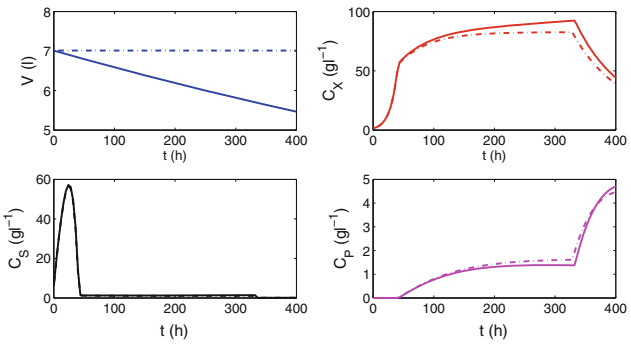


Fig. 2 Process variable profiles—comparison (*dashdot* CVol, *full* CVap)

Strategy results comparison

First, the constant vaporization (CVap) strategy has been tested on simulations with initial volume $V_0 = 7$ l and compared to the constant volume (CVol) strategy. Figure 2 shows very satisfactory cultivation results and reveals a slight superiority of the CVap strategy. It is due to the fact that the effect of input feed flow is inversely proportional to the actual amount of the broth in tank. While the effect of the input feed flow is always the same with the constant volume strategy, with the constant vaporization its positive influence improves as the volume decreases with time. Next, from the picture, it is obvious that the cultivation period that contributes to the final product concentration $C_P(t_{\mathrm{end}})$ the most takes approximately the last 75 h. A rapid product concentration increase can be observed during this period; however, the biomass concentration decreases badly. This has a simple biological explanation—as can be seen from the characters of both μ and π (see Eq. 2), increasing one of them, the second one decreases, which corresponds to the fact that either the biomass population growth or the penicillin production is being preferred at the very same time.

Volume dependency

Next, the CVol and CVap strategies have been tested on multiple simulations with various initial volume V_0. Initial volume conditions have been chosen as linearly increasing, $V_0(l) = 7 + 3k$, $k \in \{0, 1, \ldots, 10\}$.

Looking at the Fig. 3, it can be seen that with increasing V_0, final product concentration $C_P(t_{\mathrm{end}})$ decreases for both the strategies which can be (once again) explained as the

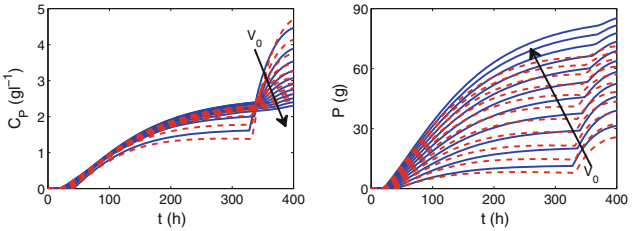


Fig. 3 Volume-dependency of C_P and P profiles (*blue full* CVol, *red dashed* CVap)

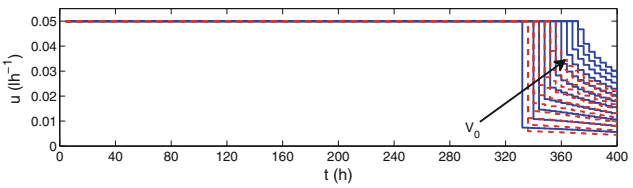


Fig. 4 Volume-dependency of input profiles (*blue full* CVol, *red dashed* CVap)

consequence of the inverse proportional effect of the actual initial volume. On the other hand, the total amount of product P increases with initial volume V_0 increase. This is due to the fact that the total product amount P is proportional not only to product concentration C_P but also to broth volume, $P = C_P V$. From this point of view, the constant volume strategy is able to obtain better results as the volume is held constant—with the constant vaporization, the volume decreases steadily and, thus, the total amount of product at t_{end} is lower than with the CVol strategy. Nevertheless, comparing the quality of the cultivation in the sense of the product concentration, CVap strategy is the more succesful candidate.

Yet, another interesting tendency is to be observed from Fig. 4—it is a convergency of input profiles to a high-saturation-valued vector with V_0 increase. From technological point of view, this is caused by the increase of V_0/U_{max} ratio—the higher the volume is, the more feed is needed to keep the whole system developing and the higher the V_0/U_{max} ratio is, the longer a high-saturated input must be applied.

To inspect the effect of various V_0 in more details, another set of simulations has been performed; however, with a constant ratio $V_0/U_{max} = 7/0.05$. The initial volume V_0 has been set linearly growing as in the previous simulation set.

Figure 5 shows that holding the V_0/U_{max} ratio fixed, C_P profiles aggravation (namely the final product concentration decay) for CVol strategy is not as drastic as in the previous case and, moreover, the C_P profile for CVap strategy does not change at all. However, this is to be expected as with fixed V_0/U_{max} ratio and the same initial concentrations, the system parameters does not change at

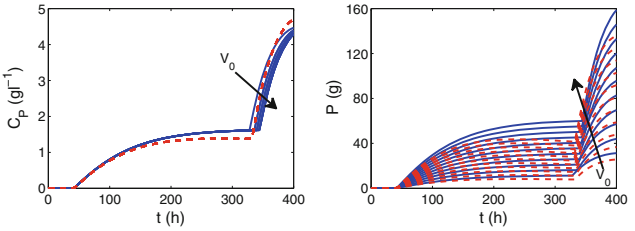


Fig. 5 Volume-dependency of C_P and P profiles, fixed V_0/U_{max} ratio (*blue full* CVol, *red dashed* CVap)

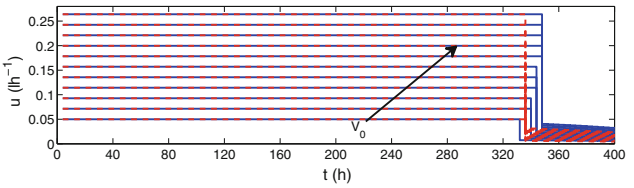


Fig. 6 Volume-dependency of input profile, fixed V_0/U_{max} ratio (*blue full* CVol, *red dashed* CVap)

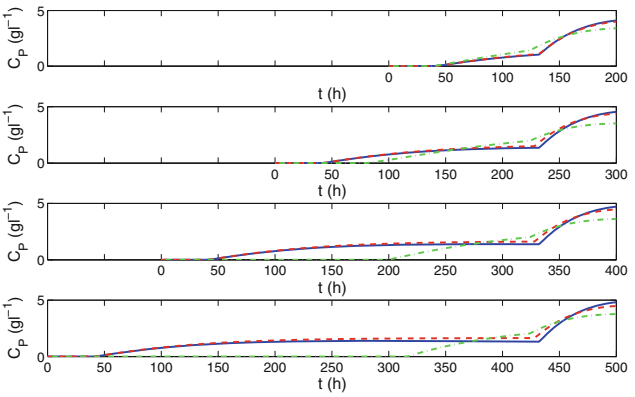


Fig. 7 C_P profiles for various cultivation lengths, comparison (*green dashdot* CG, *red dashed* CVol, *blue full* CVap

all and the system with higher V_0 is an exact scale-up of the lower V_0 one. The scale-up claim is supported by the Fig. 6, where input profiles for various V_0 are shown and it is obvious that the dynamical character of the CVap input profile remains the same and the vectors are multiplied by the V_0/U_{max} ratio.

Cultivation length dependency

As has already been mentioned, cultivation length is considered to be constant, yet it can be chosen from a set {200,300,400,500} h. Figure 7 compares cultivation with classical gradient method (presented in [16]) to the CVol and CVap strategy, respectively. For every chosen cultivation length, it is obvious that the CVap strategy achieves better results than the other ones and the product concentrations at the final time $C_P(t_{end})$ are higher.

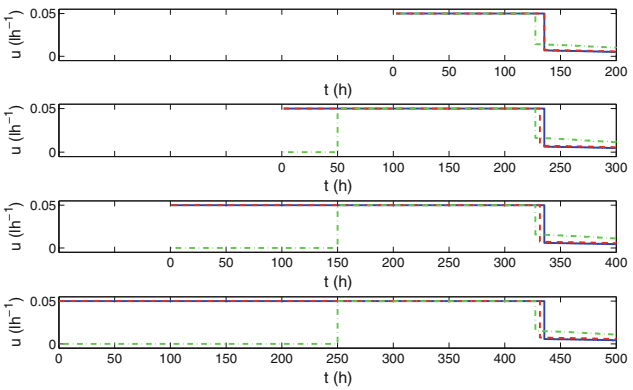


Fig. 8 Input profiles for various cultivation lengths, comparison (*green dashdot* CG, *red dashed* CVol, *blue full* CVap)

Looking at the Fig. 8, convergence of input profiles to a certain "superprofile" can be seen. Similar kind of convergence has already been mentioned in [15] as well. However, although the profiles are stable backward in time, they do not settle down at the same value (here, we assume settling down in negative march of time). In negative time, CG method settles down on a zero value, while CVol and CVap strategies obtained the input profiles settle down on upper saturation. The fact that volume is held constant (constantly decreasing, respectively) by the second virtual input (virtual due to the fact that it is not considered in optimization) and it cannot dynamically aggravate the product concentration profile enables to deliver more feed into the cultivation tank without negative effect of volume increase and, thus, with better fed biomass population, the product concentrations obtained at the end of the cultivation are higher.

Choice of gradient search step

In the previous sections, constant search step α has been assumed. Although gradient methods are able to find the (closest local) optimum, in industrial practice the quality of the solution (product concentration at final time) is of similar importance level as the promptness of the optimization algorithm (which is nothing but the convergence property). Convergence speed of the gradient method is directly related to the choice of the step with which the descent is performed—in this section, various families of gradient search steps with the effort to find the best one are investigated and evaluated. Let us remark that this evaluation can be performed and generalized for other strategies as well.

I *Fixed step family (FSF)*—the first and the simplest family contains search steps α which are constant over the whole duration of optimization. These steps are chosen as $\alpha_q = q \times 10^{-4}$, $q \in \{1, 2, \ldots, 10\}$. Fixed search step family members are then denoted as Fα_q, e.g., F1e-3 stands for fixed step approach with $\alpha = 1 \times 10^{-3}$.

II *Parabola-minimizing family (PMF)*—this family unifies approaches looking for step α_k as a minimum of parabola. The main idea is that (having computed the gradient $\partial\mathcal{H}_k/\partial v$ at the k-th iteration of the procedure described by (13)) the value of optimization criterion at this iteration $\mathcal{J}_k$ is assumed to be a quadratic function of the step α,

$$\mathcal{J}_k(\alpha) = K_{2,k}\alpha^2 + K_{1,k}\alpha + K_{0,k}. \tag{14}$$

Under this assumption, the minimum of this parabola can be found analytically choosing three different steps $[\alpha_a, \alpha_b, \alpha_c]$, computing the corresponding values $[\mathcal{J}_k|_{\alpha=\alpha_a}, \mathcal{J}_k|_{\alpha=\alpha_b}, \mathcal{J}_k|_{\alpha=\alpha_c}]$ for input vectors $v_{k-1} + \alpha_a\partial\mathcal{H}_k/\partial v$, $v_{k-1} + \alpha_b\partial\mathcal{H}_k/\partial v$ and $v_{k-1} + \alpha_c\partial\mathcal{H}_k/\partial v$ and determining coefficients $K_{0,k}$, $K_{1,k}$, $K_{2,k}$. As the parabolic approximation of the criterion $\mathcal{J}_k$ might be inaccurate in certain cases, we consider a set of triplets $[\alpha_a, \alpha_b, \alpha_c]$ at which the criterion is evaluated as follows:

$$\begin{aligned}&[\alpha_a, \alpha_b, \alpha_c] = [0 \times \alpha_{bas}, 1 \times \alpha_{bas}, 2 \times \alpha_{bas}],\\ &\alpha_{bas} = \frac{q}{2} \times 10^{-4}, q \in \{1, 2, \ldots, 8\}.\end{aligned} \tag{15}$$

Let us note that with $\alpha_a = 0$, $\mathcal{J}_k|_{\alpha=\alpha_a} = \mathcal{J}_{k-1}$. With this choice, one third of computational effort can be spared. The search step α_k which is then applied at the k-th iteration is computed as

$$\alpha_k = \arg\min(K_{2,k}\alpha^2 + K_{1,k}\alpha + K_{0,k}). \tag{16}$$

From now on, members of parabola-minimizing step family are denoted as Pmq, e.g., Pm7 represents approach where $\alpha_{bas} = \frac{7}{2} \times 10^{-4}$.

III *General-curve-minimizing family (GCMF)*—approaches performing exhaustive line search are grouped in this family. Two different sub-branches are explored: (1) a sub-branch considering linear distances between the search steps, (2) and the one performing brute-force search on logarithmically spaced vector of steps. The members of the first sub-branch are denoted as GmLin$s$ with s being the spacing of the linearly increasing vector of the examined search steps, $\alpha_{k,j} = j \times s$, $j \in \{1, 2, \ldots, j_{\max}\}$, $0 < \alpha_{k,j} \leq 1$. Here, $j_{\max}$ is the value of $1/s$ truncated to zero decimal digits,

$$j_{\max} = trunc(1/s, 0) = \lfloor 1/s \rfloor. \tag{17}$$

In most cases, a quasi-convexity of $\mathcal{J}$ as a function of α is assumed, it means that if for certain $\alpha_{k,j}$ from the explored vector $\mathcal{J}_{k,j} > \mathcal{J}_{k,j-1}$, the line search at that α-search

iteration is terminated and the applicable step for the gradient-optimization iteration is chosen as $\alpha_k = \alpha_{k,j-1}$. Such approaches are then denoted GmLin*s*QC to remark the quasi-convexity assumption, e.g., GmLin5e-4QC performs exact line search through a vector of 2,000 steps α with linear spacing $s = 5 \times 10^{-4}$ under assumption that $\mathcal{J}(\alpha)$ is quasi-convex. The second sub-branch tries to reduce the number of searched steps and it involves approaches searching through logarithmically spaced vector of steps α. Regarding this sub-branch, two approaches are considered: GmLog13 exploits vector of 13 values from 1×10^{-4} to 1 with logarithmic spacing, while GmLog13QC takes the same vector into consideration and adds quasi-convexity assumption.

Results with enhanced search step choice

In this section, results of the three-step choice families described in the previous section are presented.

Table 3 brings comparison of optimality $\mathcal{J}(\mathrm{gl}^{-1})$, number of iterations-to-converge ItC (−) and time-to-converge TtC (min) for the inspected families and their member approaches. "NA" value means that the convergence has not been achieved.

Table 3 Results and computational demands comparison

Family	Member	$\mathcal{J}$	ItC	TtC
FSF	F1e-4	4.71	48,287	707
	F2e-4	4.71	24,137	336
	F3e-4	4.71	16,092	224
	F4e-4	4.71	12,070	167
	F5e-4	4.71	9,651	133
	F6e-4	4.71	8,040	110
	F7e-4	4.71	6,890	94
	F8e-4	NA	NA	NA
	F9e-4	NA	NA	NA
	F1e-3	NA	NA	NA
PMF	Pm1	4.70	275	7
	Pm2	4.69	366	10
	Pm3	4.70	428	11
	Pm4	4.70	453	12
	Pm5	4.70	470	12
	Pm6	4.70	726	19
	Pm7	NA	NA	NA
	Pm8	NA	NA	NA
GCMF	GmLin2e-4QC	4.74	886	121
	GmLin3e-4QC	4.76	2,055	108
	GmLin4e-4QC	4.76	1,187	73
	GmLin5e-4QC	NA	NA	NA
	GmLin4e-4	4.76	1,016	16,185
	GmLog13QC	4.76	1,410	52
	GmLog13	4.76	1,297	124

Regarding the FSF, it can be seen that with increase of the search step, ItC value decreases and so does TtC. This can be expected as with greater steps α, one can await faster convergence rate as the gradient method moves quicker towards the supposed minimum. A situation which often occurs when using gradient method can be observed for steps $\alpha \geq 8 \times 10^{-4}$—from this value, very large search steps destroy the convergence properties of gradient search which is known to be susceptible to the oversized step choice. Here it could be noted that fixed step family is the most computationally demanding from the three search steps families with sovereignly highest ItC values. This is the price to be paid for the fact that the gradient search with sufficiently small search step guarantees convergence to the closest local minimum. Its local-minimum-convergence is the next disadvantage—as can be seen later, this can be overcome exploring larger part of step-space.

On the other hand, the PMF converges usually extremely fast compared to the other families. However, with increasing distance between the examined α_a, α_b and α_c, the approximation is less and less accurate which reflects in ItC increase and for very large α_{bas}, the converge is not reached.

Looking at the $\mathcal{J}$-column, GCMF achieves the most superior results. Also this family demonstrates that larger steps bring faster convergence to a small environment of the minimum but with very large search steps, convergence is not guaranteed. An interesting insight offers comparison of GmLin4e-4QC and GmLin4e-4 approaches—as they both achieve the same value of the final product concentration, the quasi-convexity assumption is proved right. Yet, the difference between TtC values is enormous. Although GmLin4e-4 spares approximately 150 iterations, it is clear that the most of the α-search iterations performed at every optimization iteration k are redundant. The same is supported by the approaches using logarithmically-spaced α-vectors. Moreover (as is shown in Fig. 9, where the steps α_k which are finally applied at particular iteration k are depicted), the steps which are most attractive from the convergence point of view are quite small and large steps

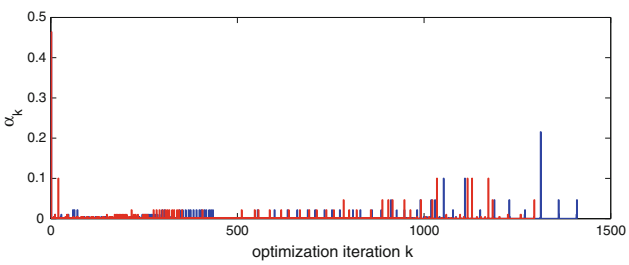


Fig. 9 α_k evolution (*blue* GmLog13QC, *red* GmLog13)

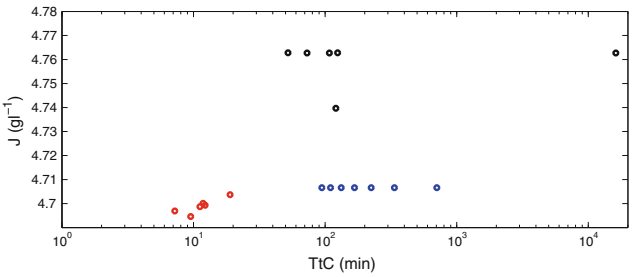


Fig. 10 Step choice families comparison (*blue* FSF, *red* PMF, *black* GCMF)

α_k occur rather rarely. This is the consequence of nonlinear dynamics of the system which turns the optimization task into a nonconvex one and the υ-gradient of the Hamiltonian $\partial\mathcal{H}/\partial\upsilon$ which is used in the best direction search is usually only locally valid. Therefore, it is more convenient to distribute the examined search steps with higher intensity tightly around the current point in the explored optimization variables space while including a few outliers, the convergence gets faster.

Comparison of $\mathcal{J}$ and TtC for the successful members of particular step choice families can be seen in Fig. 10.

Conclusion

This paper follows the patterns suggested in previous publications of the authors—a model of the controlled system involving the second input variable is derived, a neat way of problem reduction to a single-input optimization is performed and thanks to this, a successful constant vaporization control strategy is introduced. The necessary model adaptation is performed so that it comports with the strategy requirements. Results are verified on a set of numerical experiments and discussed in detail.

The comparison obtained by verifying the constant vaporization strategy on a set of numerical simulations and confronting it with the previously introduced methods can be summarized as very encouraging—CVap strategy achieves better results which suggest its possible industrial use. A "superprofile" convergence (observed in earlier publications) occurs in this case as well, and this supports the claim that it is a property of the optimization issues where the fixed time, fixed initial condition and free terminal condition are considered.

Next, various ways of gradient search step choice are suggested, explained and compared. The most time-sparing group of approaches appears to be the PMF which approximates the cost criterion by a parabola while regarding the final product concentration value, the most successful is the GCMF. As a tradeoff between optimality and time consumption, logarithmically spaced vector of search steps can be chosen—it combines fast convergence rate and higher obtained final product concentration and adding the quasi-convexity assumption, it seems to be the fair choice for further utilization within the optimization routine.

Acknowledgment This research has been supported by the Czech Science Foundation through the grant no. 13-20433S.

References

1. Lee S (1996) High cell-density culture of *Escherichia coli*. Trends biotechnol 14(3):98–105
2. Gregory M, Turner C (1993) Open-loop control of specific growth rate in fed-batch cultures of recombinant *E. coli*. Biotechnol Techn 7(12):889–894
3. Suzuki T, Yamane T, Shimizu S (1990) Phenomenological background and some preliminary trials of automated substrate supply in pH-stat modal fed-batch culture using a setpoint of high limit. J Ferment Bioengineer 69(5):292–297
4. Siimes T, Linko P, von Numers C, Nakajima M, Endo I (1995) Real-time fuzzy-knowledge-based control of Baker's yeast production. Biotechnol bioengineer 45(2):135–143
5. Horiuchi J (2002) Fuzzy modeling and control of biological processes. J biosci bioengineer 94(6):574–578
6. Zhu G, Zamamiri A, Henson M, Hjortsø M (2000) Model predictive control of continuous yeast bioreactors using cell population balance models. Chem Engineer Sci 55(24):6155–6167
7. Azimzadeh F, Galan O, Romagnoli J (2001) On-line optimal trajectory control for a fermentation process using multi-linear models. Comput Chem Engineer 25(1):15–26
8. Alford J (2006) Bioprocess control: Advances and challenges. Comput Chem Engineer 30(10):1464–1475
9. Nagy Z (2007) Model based control of a yeast fermentation bioreactor using optimally designed artificial neural networks. Chem Engineer J 127(1):95–109
10. Alvarez M, Stocks S, Jørgensen S (2009) Bioprocess modelling for learning model predictive control (l-mpc). In: Computational intelligence techniques for bioprocess modelling, supervision and control, pp 237–280
11. Zhang H, Lennox B (2004) Integrated condition monitoring and control of fed-batch fermentation processes. J Process Contr 14(1): 41–50
12. Ashoori A, Moshiri B, Khaki-Sedigh A, Bakhtiari M (2009) Optimal control of a nonlinear fed-batch fermentation process using model predictive approach. J Process Contr 19(7):1162–1173
13. Dai Y, Yuan Y (2000) A nonlinear conjugate gradient method with a strong global convergence property. SIAM J Optim 10(1):177–182
14. Birgin E, Martínez J (2001) A spectral conjugate gradient method for unconstrained optimization. Appl Math optim 43(2):117–128
15. Čelikovský S, Papáček Š, Cervantes-Herrera A, Ruiz-Leon J (2010) Singular perturbation based solution to optimal microalgal growth problem and its infinite time horizon analysis. IEEE Trans Automat Contr 55(3):767–772
16. Pčolka M, Čelikovský S (2012) Gradient method optimization of penicillin production. In: 24th Chinese Control and Decision Conference (CCDC), 2012. IEEE, pp 74–79
17. Pčolka M, Čelikovský S (2012) Gradient method optimization of penicillin production: New strategies. In: 20th Mediterranean Conference on Control & Automation (MED), 2012. IEEE, pp 1235–1240
18. Pčolka M, Čelikovský S (2012) Multiple-Input Cultivation Model Based Optimization of Penicillin Production. In: 51st IEEE Conference on Decision and Control, Hawaii, USA

19. Van Impe J, Bastin G (2002) Optimal adaptive control of fed-batch fermentation processes with multiple substrates. In: Control Applications, 1993., Second IEEE Conference on. IEEE, pp 469–474
20. Constantinides A, Spencer J, Gaden Jr E (1970) Optimization of batch fermentation processes. II. Optimum temperature profiles for batch penicillin fermentations. Biotechnol Bioengineer 12(6): 1081–1098
21. Contois D (1959) Kinetics of bacterial growth: relationship between population density and specific growth rate of continuous cultures. Microbiology 21(1):40
22. Briggs G, Haldane J (1925) A note on the kinetics of enzyme action. Biochem j 19(2):338
23. Álvarez L, García J, Urrego D Control of a fedbatch bioprocess using nonlinear model predictive control.
24. Azimzadeh F, Galán O, Barford J, Romagnoli J (1999) On-line optimization control for a time-varying process using multiple models: A laboratory scale fermenter application. Comput Chem Engineer 23:S235–S239
25. Foss B, Johansen T, Sørensen A (1995) Nonlinear predictive control using local models–applied to a batch fermentation process. Contr Engineer Practice 3(3):389–396
26. Bryson A, Ho Y (1969) Applied optimal control. Blaisdell, New York