## Invited Paper

# Intelligent Optimization of a Mixed Culture Cultivation Process

**Petia Koprinkova-Hristova[1*], Georgi Kostov[2], Silviya Popova[3]**

[1]*Institute of Information and Communication Technologies*
*Bulgarian Academy of Sciences*
*Acad. G. Bonchev Str., Bl. 25A, 1113 Sofia, Bulgaria*
*E-mail: pkoprinkova@bas.bg*

[2]*Department Technology of Wine and Brewing*
*University of Food Technologies*
*26 Maritza Blvd., 4002 Plovdiv, Bulgaria*
*E-mail: george_kostov2@abv.bg*

[3]*Institute of System Engineering and Robotics*
*Bulgarian Academy of Sciences*
*Acad. G. Bonchev Str., Bl. 2, 1113 Sofia, Bulgaria*
*E-mail: popova_silvia2000@yahoo.com*

[*]*Corresponding author*

**Received: July 07, 2014** **Accepted: December 22, 2014**

**Published: April 20, 2015**

***Abstract:*** *In the present paper a neural network approach called "Adaptive Critic Design" (ACD) was applied to optimal tuning of set point controllers of the three main substrates (sugar, nitrogen source and dissolved oxygen) for PHB production process. For approximation of the critic and the controllers a special kind of recurrent neural networks called Echo state networks (ESN) were used. Their structure allows fast training that will be of crucial importance in on-line applications. The critic network is trained to minimize the temporal difference error using Recursive Least Squares method. Two approaches – gradient and heuristic – were exploited for training of the controllers. The comparison is made with respect to achieved improvement of the utility function subject of optimization as well as with known expert strategy for control the PHB production process.*

***Keywords:*** *Mixed culture cultivation, PHB production process, Adaptive critic design (ACD), Optimization, Echo state networks (ESN).*

## Introduction

The mixed culture systems are quite common in nature: the human body, waste water treatment, ecosystems are some of well known examples. In such systems one microorganism assimilates substrate A and converts it to metabolite B which is converted by another microorganism to metabolite C. Since the change in culture conditions affects all microorganisms differently it is difficult to control such processes in an optimal way.

The present paper considers the mixed culture system where sugars (glucose) were converted to lactate by the microorganism *L. delbrueckii* and then the lactate was converted to PHB (poly-β-hydroxybutyrate) by the microorganism *R. euthropha*. The main product – PHB –

is biodegradable polymer used as thermoplastic in food and drug industry. Hence the main purpose of the process control strategy is to maximize the outcome of this product accounting for the needs and mutual relations of both microorganisms in the culture.

By now there are known several approaches to this problem. In [8, 28] different control strategies were exploited separately or in combination: to maintain the lactate concentration at an given optimal level using dissolved oxygen concentration as control variable, to maintain the glucose concentration at a given optimal level by its feeding rate, to change the set point of the glucose concentration according to the lactate concentration deviation from its set point. In [5] it was proposed to monitor the lactate production and consumption rates in order to determine the needs of the two microorganisms and depending on them to feed glucose or to change dissolved oxygen concentration. Another approach is adaptive control strategy proposed in [21] that determines the optimal glucose feeding rate based on the known from [28] optimal level of the lactate concentration or glucose concentration and monitoring of the second microorganism's concentration and lactate [20]. In [4] it was proposed to maximize the process productivity by controlling the mixing intensity. In [19] an intelligent approach to optimization of the glucose and ammonium time profiles is proposed. It uses neural networks for process model and feed-back controller. In [9] fuzzy control approach is proposed that combines the expert knowledge about the lactate concentration dependence on the set points of dissolved oxygen and glucose concentrations.

In previous work [10] the neural network approach called "Adaptive Critic Design" (ACD) was applied to synthesis of sugar's concentration optimal time profile for the process. In [12] the same approach was extended to synthesis of optimal time profiles of all three main substrates (sugar, nitrogen source and dissolved oxygen). However by far the control scheme was open loop, i.e. there was no feedback from the process state to the controllers. Instead only time profiles of the set points of the control variables were adjusted. In [16] the ACD was extended to the closed-loop version of ACD using ESN structure for the controllers and two algorithms (gradient and heuristic) for training of these controllers accounting for ESN structure peculiarities were developed.

Here we focus on the optimization results achieved on the task of PHB production maximization. The comparison of both training algorithms developed in [16] is made with respect to achieved improvement of the utility function subject of optimization as well as with expert strategy for control of the PHB production process.

The rest of the paper is organized as follows: in Section 2 the PHB production process is described and the model used for its simulation is presented; next the basics of ACD optimization technique is presented; Section 2.3 gives brief description of the ESN structures is given; section three presents the optimization task that was solved – maximization of product outcome of the PHB process – and the results are presented and discussed; the paper finishes with the conclusion section.

## Problem statement

### *PHB production process*

In [28] the PHB production process was modeled by a system with six nonlinear ordinary equations as follows:

$$\frac{dX_1}{dt} = \mu_1(S, P, DO)X_1 - \frac{F_S}{V}X_1 - \frac{F_n}{V}X_1 \tag{1}$$

$$\frac{dS}{dt} = -\nu_1(S,P,DO)X_1 + \frac{F_S(S_F - S)}{V} - \frac{F_n}{V}S \tag{2}$$

$$\frac{dP}{dt} = \sigma_1(S,P,DO)X_1 - \nu_2(P,DO,N)X_2 - \frac{F_S}{V}P - \frac{F_n}{V}P \tag{3}$$

$$\frac{dX_2}{dt} = \mu_2(P,DO,N)X_2 - \frac{F_S}{V}X_2 - \frac{F_n}{V}X_2 \tag{4}$$

$$\frac{dN}{dt} = -\nu_3(P,DO,N)X_2 + \frac{F_n(N - N_F)}{V} - \frac{F_S}{V}N \tag{5}$$

$$\frac{dQ}{dt} = \sigma_2(N)X_2 - \frac{F_S}{V}Q - \frac{F_n}{V}Q \tag{6}$$

$$\frac{dV}{dt} = F_S + F_n \tag{7}$$

Here the main process state variables are: $X_1$ – cell concentration of *L. delbrueckii*; $X_2$ – cell concentration of *R. euthropha*; $S$ – glucose concentration; $P$ – lactate concentration; $N$ – $NH_3$ concentration, and $Q$ – product, i.e. PHB concentration. The bioreactor's volume is denoted by $V$. The feeding rates of the two main substrates (sugar and nitrogen source) are denoted by $F_S$ and $F_n$, respectively. The specific growth and consumption rates are described as follows:

$$\mu_1(S,P,DO) = \frac{\mu_{m1}(DO)}{Y_S + S}\left(1 - \frac{P}{P_m}\right)^n \tag{8}$$

$$\nu_1(S,P,DO) = \frac{\sigma_1(S,P,DO)}{Y_{P/S}(DO)} \tag{9}$$

$$\sigma_1(S,P,DO) = \alpha\mu_1(S,P,DO) + \beta(S,DO) \tag{10}$$

$$\beta(S,DO) = \frac{\beta_m(DO)S}{K_S + S} \tag{11}$$

$$\nu_2(P,DO,N) = \frac{\mu_2(P,DO,N)}{Y_{X_2/P}(DO)} \tag{12}$$

$$\mu_2(P,DO,N) = \left(\frac{\mu_{m2}(DO)P}{K_P + P + P^2/K_i}\right)\left(\frac{N}{K_N + N}\right) \tag{13}$$

$$\nu_3(P,DO,N) = \frac{\mu_2(P,DO,N)}{Y_{X_2/N}(DO)} \tag{14}$$

$$\sigma_2(N) = q_m\left(\frac{k_N}{k_N + N}\right) \tag{15}$$

The main parameters' dependence on dissolved oxygen concentration ($DO$) is as follows:

$$\mu_{m1}(DO) = a_1 e^{-a_2 DO} + a_3 \tag{16}$$

$$Y_{P/S}(DO) = b_1 e^{-b_2 DO} + b_3 \tag{17}$$

$$\beta_m(DO) = c_1 e^{-c_2 DO} + c_3 \tag{18}$$

$$\mu_{m2}(DO) = d_1 e^{-d_2 DO} + d_3 \tag{19}$$

$$Y_{X_2/P}\left(DO\right) = f_1 e^{-f_2 DO} + f_3 \tag{20}$$

In our simulation $Y_{X_2/N} = const$. The temperature is thermostated at 37 °C. The pH was maintained at the specified value by adding NaOH or NCl solution. The dissolved oxygen concentration was maintained at the set point by changing the agitation speed and/or air flow rate.

## *ACD approach*

ACD [22] originate from one side as a method approximating Bellman's dynamic programming [2, 3] and from the other side as gradient version of associative "learning from experience" called Reinforcement Learning (RL) [1]. During the last thirty years theoretical developments in this field led to numerous variations of optimal control approaches [17]. The core of the methods is approximation of Bellman's equation via neural network called "heuristic adaptive critic". Training of a critic is done minimizing temporal difference (TD) error [27] thereby mimicking the brain's ability to learn how to predict future outcomes on the basis of previous experience without awaiting the final results from future actions. The key component of ACD training and solving the optimization task is the backpropagation method that is gradient algorithm based on the chain rule of derivative calculation [29]. In contrast, the RL uses Hebbian or associative learning law for both critic and controller (called actor in terms of RL) networks. Usually the critic is trained off-line since it needs a collection of a variety of data from the beginning to the end of several process runs. Combination between off-line and on-line learning is also considered [23]. True on-line applications of ACD approaches, however, needs very fast training algorithms [24]. In highly non-linear environments the necessity for additional feedback connections arises, which further complicates the on-line training. In such cases the application of backpropagation trough time (BPTT) [29] is an alternative. However, it is impossible to be used in an on-line mode. Instead of that the Extended Kalman Filter (EKF) method [7] is usually applied, which is more complicated and resource demanding. Hence it is crucial to work towards finding simply trainable recurrent network structures for ACD schemes.

ACD approach was already applied for optimization of biotechnological processes [6], however in off-line mode.

In search of fast trainable neural network architectures in [12-14] it was proposed to use recently developed class of Recurrent Neural Networks (RNNs) called Echo State Networks (ESNs) [7]. Their structure incorporates a dynamic reservoir of neurons that is generated randomly and a fast trainable readout layer. These allow on-line adaptation via Recursive Least Square (RLS) method [7] as well as calculation of needed derivatives with much less computational effort [14].

From biological point of view however, the gradient learning is considered as non-plausible. It is claimed that associative learning algorithms like Hebbian law are closer to the biological neurons behavior. That is why in [15] it was proposed to incorporate associative learning laws within ACD scheme. That was done via training of actor with associative manner like in [1]. However in [15] the actor was not entire network structure but only time profile of the control variable. In [16] the approach was extended to the colsed loop scheme with ESN structures for the actors (controllers) and the gradient and associative training algorithms were extended accounting for ESN structure peculiarities.

The main scheme of the on-line approach according to [26] is given on Fig. 1 below.

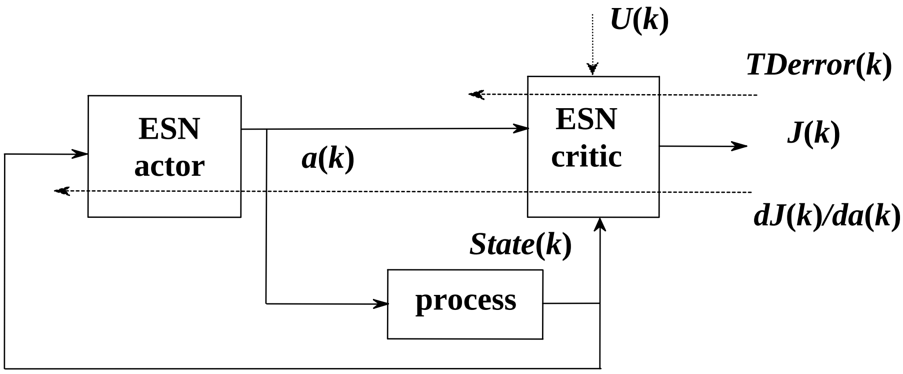


Fig. 1 Adaptive Critic Design optimization approach

Here the dashed lines represent the training cycle. The vector *State*($t$) represents the object state vector, $a(t)$ is control (action) variable. The critic NN has to be trained to predict the utility function $U(t)$ by approximating Bellman's equation as follows:

$$J\left(State(k),\, a(k)\right) = \sum_{t=0}^{k} \gamma^{t} U\left(State(t),\, a(t)\right) \tag{21}$$

where $\gamma$ is a discount factor.

The critic network is trained so as to minimize the TD error:

$$TDerror(k) = J(k) - U(k) - \gamma J(k+1) \tag{22}$$

The action ESN represents the controller that has to be adjusted so as generated by its control actions maximize (minimize) the utility function. The feedback connection from the process state to the controller can include the full state vector or some of state variables. The dashed lines represent the training cycles of critic and actor respectively.

Concerning the action ESN, it has to be trained so as to generate proper control actions. In the classical ACD it is done via backpropagation of utility [29] that is gradient descent training. Here it is compared with biologically plausible associative learning algorithm adopted from early RL scheme [1]. Both algorithms are described in details in [16].

## *Echo state networks*

ESNs are a kind of recurrent neural networks that arise from so called "reservoir computing approaches" [18]. The basic ESN structure is shown in Fig. 2 below.

The ESN output vector denoted here by *out*($k$) (it will be $J(k)$ or $a(k)$ for critic and action networks respectively) for the current time instance $k$ is usually a linear function of its input and current state:

$$out(k) = f^{out}\left(W^{out}\left[in(k), R(k)\right]\right) \tag{23}$$

Here, $in(k)$ is a vector of network inputs and $R(k)$ a vector composed of the reservoir neuron states; $f^{out}$ is a linear function (usually the identity), $W^{out}$ is a noutx($n_{in} + n_R$) trainable matrix (here $n_{out}$, $n_{in}$ and $n_R$ are the sizes of the corresponding vectors *out*, *in* and *R*).

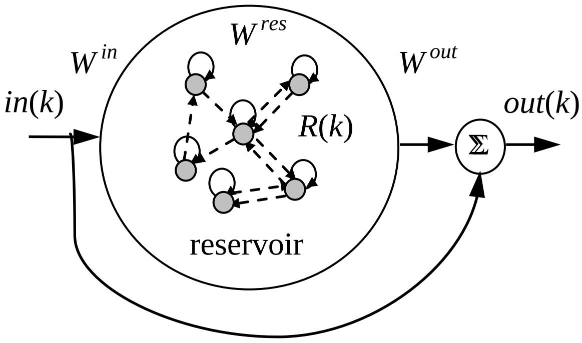


Fig. 2 Echo state networks structure

The neurons in the reservoir have a simple sigmoid output function $f^{res}$ (usually hyperbolic tangent) that depends on both the ESN input $in(k)$ and the previous reservoir state $R(k-1)$:

$$R(k) = f^{res}\left(W^{in} in(k) + W^{res} R(k-1)\right) \tag{24}$$

Here $W^{in}$ and $W^{res}$ are $n_{in} \times n_R$ and $n_R \times n_R$ matrices that are randomly generated and are not trainable. There are different approaches for reservoir parameter production [18]. A recent approach used in the present investigation is proposed in [25]. It is called intrinsic plasticity (IP) and suggests initial adjustment of these matrices, aiming at increasing the entropy of the reservoir neurons outputs. For on-line training, the RLS algorithm [7] was used.

## Results and discussion

The main goal of the optimization is to maximize the outcome of the process product $Q$. Hence the utility function at each time step $k$ will be:

$$U(k) = Q(k)V(k) \tag{25}$$

and the overall utility for the process with $N$ time steps will be:

$$U_{sum} = \sum_{k=0}^{N} U(k) \tag{26}$$

For the PHB production process the vector $State(k)$ (from Fig. 1) includes all main process state variables, i.e.:

$$State(k) = \left[X_1(k), S(k), P(k), X_2(k), N(k), Q(k)\right] \tag{27}$$

The ideology of the process control scheme is described in more details in [11, 28]. We suppose that all concentration controllers work properly and that they are able to follow the set points. Hence the optimization task to be solved is to determine the proper values of these set points at each moment. Hence the control vector consists of the three set points as follows:

$$a(k)=\left[S^*(k),\ N^*(k),\ DO^*(k)\right] \tag{28}$$

Following the ACD scheme from Fig. 1, for each control variable a corresponding ESN action network was trained using both gradient and associative rules described in [16]. In present work we choose to have only one input of each action ESN – the key intermediate metabolite $P$ since it is on-line measurable and its concentration is of crucial importance for process trend.

All control variables have imposed restrictions in terms of minimum and maximum values allowed. They were included in the utility function as follows:

$$U(k)=Q(k)V(k)-\frac{1}{2}ra^2(k) \tag{29}$$

Here $ra(k)$ is a kind of "punishment" signal in the case when calculated by ESN control action is outside allowed interval [$a_{min}$, $a_{max}$] as follows:

$$ra(k)=\begin{cases}-\left(a_{min}-a(k)\right), & a(k)<a_{min}\\ \left(a(k)-a_{max}\right), & a(k)>a_{max}\end{cases} \tag{30}$$

For the ESN critic training and simulation a Matlab toolbox from [7] with our improvements for IP training as in [25] was used. The critic network has 9 inputs (6 for the process state variables plus 3 for the control actions), 10 reservoir neurons and 1 output. The action networks have one input, one output and 5 neurons in the reservoir each. All reservoir neurons have hyperbolic tangent output function. The initial set point profiles were taken from [11]. Detailed optimization algorithm can be found in [12]. It consists of consecutive critic and actor training iterations. Here for comparative purpose simple gradient algorithm without any improvement (such as momentum term or variable speed) was used. After every cycle of a critic plus an action training iteration parameter $\gamma$ is slightly increased until it become equal to 0.5. During first 1000 iterations $\gamma$ reaches its maximal value and within the rest of 200 iterations it was constant. Since in previous work [15] it was observed that the procedure is too sensitive to big changes in discount factor, here a small step of 0.001 was used.

In [16] it was observed that although at the beginning the gradient algorithm looks faster and it reaches bigger utility values in comparison with the associative one, by the end of iterations associative algorithm gives bigger outcome in comparison with the gradient one. Looking at convergence speed it seems almost the same for both algorithms especially after discount factor reaches its maximum value. Further improvement of both algorithms could be achieved by using variable learning rate that will allow preventing observed now big variations of the utility values during iterations. In both cases the trained actors have stable work. There was not observed uncontrolled increase of trained weights – a problem that was observed in the case of RLS training procedure before.

Fig. 3 presents time trends of all process state variables obtained with different control policies: the defined by experts strategy [28], starting point of both optimization algorithms and achieved after optimization using gradient and heuristic training of ESN controllers. On Fig. 4 the corresponding time profiles of the three control variables are presented. Fig. 5 presents achieved in these four cases utility by the end of the process.

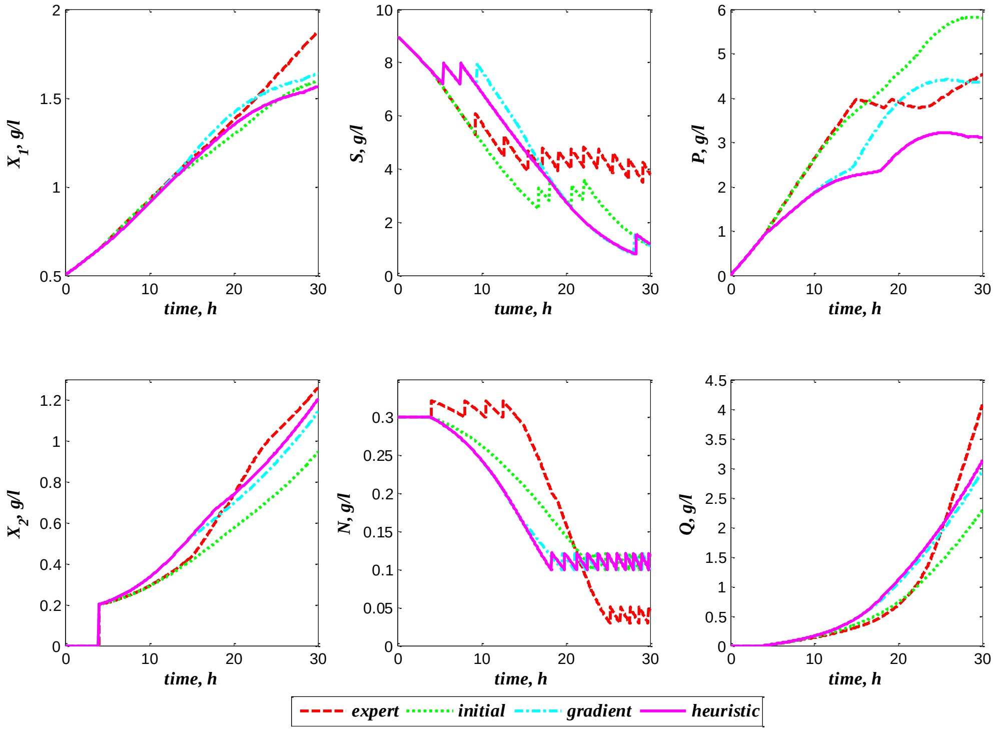


Fig. 3 Time trends of the process state variables

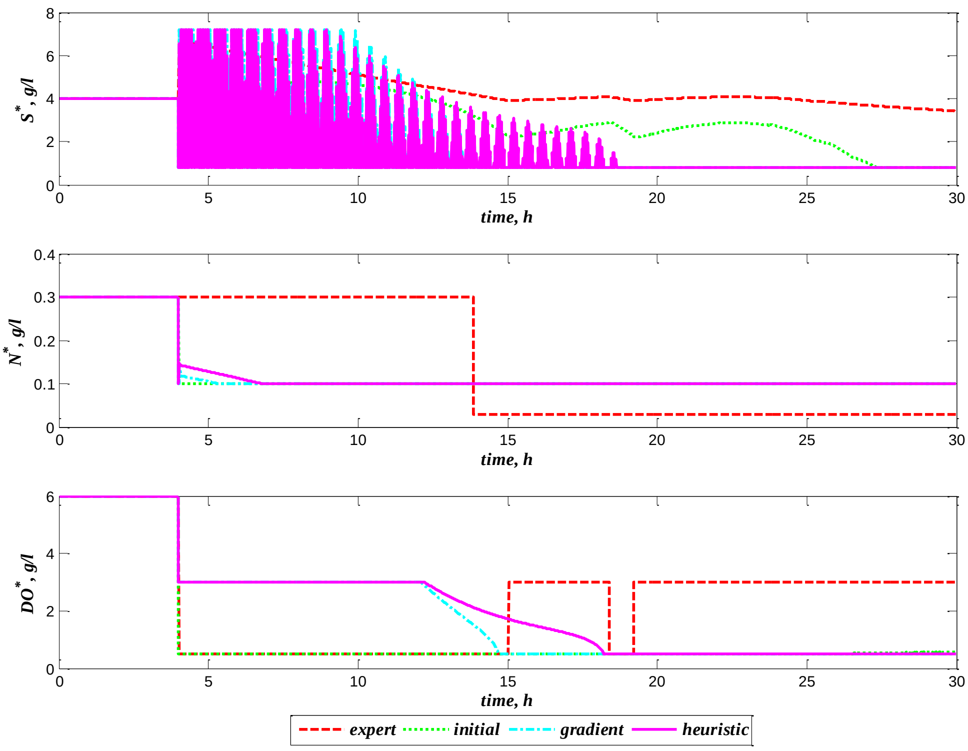


Fig. 4 Obtained time profiles of the control variables

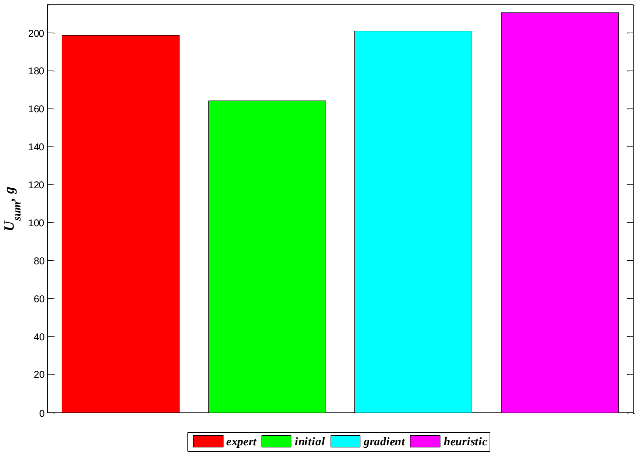


Fig. 5 Achieved utility value with four compared approaches

Looking at Fig. 3 we observe that the key intermediate product (lactate) is closer to its optimal value determined by experts in the case of gradient optimization approach while in the case of heuristic optimization its final value is lower. This led to lower concentrations of the first microorganism while the concentration of the second reaches approximately the same level as in the case of expert control approach. This effect can be explained by the lower sugar concentrations maintained by both optimization strategies in comparison to the expert strategy. However we observe that both optimization strategies are able to maintain approximately constant value of lactate at the second stage of the process than could allow longer cultivation time. In contrast the expert strategy led to undesired increase of lactate by the end of process that could suppress growth of both strains in future.

Both optimization control policies maintain higher nitrogen source concentrations at the second stage of the process. This can explain the fact that although the final PHB outcome is lower in the case of optimization strategies, its increase starts earlier than in the case of expert control and the concentration of PHB stays higher until about 25-th hour of cultivation when the expert strategy becomes more favorable.

Looking at Fig. 4, we observe that while expert strategy prescribes gradual decrease of sugar concentration after inoculation of the second strain at 4-th hour of cultivation, the optimization strategies try to maintain this concentration higher longer. However after 18-th hour both gradient and heuristic optimizations prescribe lowest possible sugar concentration to be maintained.

Concerning the nitrogen concentration, both optimization decisions are to decrease it almost immediately after inoculation of the second strain but they maintain it at higher value during the rest of the process in comparison with the expert strategy. In contrast the expert strategy prescribes to keep nitrogen source at its maximal value during 14-th hour of cultivation and then to keep it as low as possible.

While the experts prescribe to increase from time to keep dissolved oxygen concentration low after inoculation of the second strain and to increase it only from time to time during second stage of the process, both optimization strategies prescribe gradual decrease of dissolved oxygen in the broth reaching its minimal value at 18-th hour.

Observed accumulated utility during overall process (Fig. 5) shows that both optimization strategies give increased productivity in comparison to their starting point that is slightly higher than in the case of expert control strategy. We also observed that heuristic approach outperforms the gradient algorithm achieving highest total outcome of the process.

## Conclusion

The carried out simulation investigations showed that both gradient and associative learning algorithms fit well to training of the ESN structure having the role of the controllers within ACD scheme. Although results did not show significant differences in the achieved results, associative learning can be consider as the better algorithm due to the following reasons: first it is more biologically plausible and second it seems that during the iterations it showed slightly smaller variations of the utility function and better convergence characteristics in comparison with the gradient algorithm.

Concerning the obtained results for the PHB production process, both optimization strategies showed comparative performance with the expert control strategy. The main difference is in prescribed higher concentrations of the nitrogen source during second stage of the process and maintained lower glucose concentrations during this stage. It seems that these approaches will allow longer cultivation with possible better outcome at the end of the process. The investigations in this direction will be our next aim.

## Acknowledgements

*The research work reported in the paper is partly supported by the project AComIn, grant 316087, funded by the FP7 Capacity Programme (Research Potential of Convergence Regions).*

## References

1. Barto A. G., R. S. Sutton, C. W. Anderson (1983). Neuronlike Adaptive Elements that Can Solve Difficult Learning Control Problems, IEEE Transactions on Systems, Man and Cybernetics, 13(5), 834-846.
2. Bellman R. E. (1957). Dynamic Programming, Princeton, NJ: Princeton Univ. Press.
3. Bertsekas D. P., J. N. Tsitsiklis (1996). Neuro-dymanic Programming, Athena Scientific, Belmont, MA.
4. Ganduri V. S. R. K., S. Ghosh, P. R. Patnaik (2005). Mixing Control as a Device to Increase PHB Production in Batch Fermentations with Co-cultures of *Lactobacillus delbrueckii* and *Ralstonia eutropha*, Process Biochemistry, 40, 257-264.
5. Ignatova M., V. Lyubenova (2007). Adaptive Control of Fed-batch Processes for poly-β-hydroxybutyrate Production by Mixed Culture, Comptes rendus de l' Academie bulgare des Sciences, 60(5), 517-524.
6. Iyer M., D. Wunsch II (2001). Dynamic Re-optimization of a Fed-batch Fermentor Using Adaptive Critic Designs, IEEE Transactions on Neural Networks, 12(6), 1433-1444.
7. Jaeger H. (2002). Tutorial on Training Recurrent Neural Networks, Covering BPPT, RTRL, EKF and the "Echo State Network" Approach, GMD Report 159, German National Research Center for Information Technology.
8. Koprinkova-Hristova P., T. Patarinska (2006). Neural Network Modeling of the Mixed Culture Cultivation for PHB Production Process, Proceeding of 20$^{th}$ International Conference SAER'2006, September 23-24, Varna, Bulgaria, 110-117.
9. Koprinkova-Hristova P. (2007). Fuzzy Control Approach to Mixed Culture Cultivation for PHB Production Process, Proceeding of the International Conference Automatics and Informatics'07, October 3-6, Sofia, Bulgaria, III-65-III-68.

10. Koprinkova-Hristova P. (2008). ACD Approach to Optimal Control of Mixed Culture Cultivation for PHB Production Process – Sugar's Time Profile Synthesis, Proceeding IEEE Intelligent Systems IS'08, Methodology, Models, Applications and Emerging Technologies, September 6-8, Varna, Bulgaria, II, 12-29-12-32.
11. Koprinkova-Hristova P. (2008). Knowledge-based Approach to Control of Mixed Culture Cultivation for PHB Production Process, Biotechnology and Biotechnological Equipment, 22(4), 964-967.
12. Koprinkova-Hristova P., G. Palm (2010). Adaptive Critic Design with ESN Critic for Bioprocess Optimization, Lecture Notes in Computer Science, 6353, 438-447.
13. Koprinkova-Hristova P., M. Oubbati, G. Palm (2010). Adaptive Critic Design with Echo State Network, Proceeding of 2010 IEEE International Conference on Systems, Man and Cybernetics, October 10-13, Istanbul, Turkey, 1010-1015.
14. Koprinkova-Hristova P., M. Oubbati, G. Palm (2013). Heuristic Dynamic Programming Using Echo State Network as Online Trainable Adaptive Critic, International Journal of Adaptive Control and Signal Processing, 27(10), 902-914.
15. Koprinkova-Hristova P. (2014). Adaptive Critic Design and Heuristic Search for Optimization, Lecture Notes in Computer Science, 8353, 248-255.
16. Koprinkova-Hristova P. (2015). Hebbian versus Gradient Training of ESN Actors in Closed-loop ACD, Lecture Notes in Computer Science, 8962, 95-102.
17. Lenardis G. G. (2009). A Retrospective on Adaptive Dynamic Programming for Control, Proceeding of the Joint Conference on Neural Networks, Atlanta, GA, USA, June 14-19, 1750-1757.
18. Lukosevicius M., H. Jaeger (2009). Reservoir Computing Approaches to Recurrent Neural Network Training, Computer Science Review, 3, 127-149.
19. Patnaik P. R. (2005). Neural Network Designs for poly-b-hydroxybutyrate Production Optimization under Simulated Industrial Conditions, Biotechnology Letters, 27, 409-415.
20. Popova S. (2006). On-line State and Parameters Estimation based on Measurements of the Glucose in Mixed Culture System, Biotechnology and Biotechnological Equipment, 20(3), 208-214.
21. Popova S. (2007). Adaptive Control for PHB Production, Acta Universitasis Cibernesis, Series E Food Technology, XI, 17-25.
22. Prokhorov D. V. (1997). Adaptive Critic Designs and Their Applications, Ph.D. Thesis, Department of Electrical Engineering, Texas Tech University.
23. Prokhorov D. (2007). Toward Effective Combination of Off-line and On-line Training in ADP Framework, Proceeding of the IEEE Symposium on Approximate Dynamic Programming and Reinforcement Learning, ADPRL'2007, 268-271.
24. Prokhorov D., (2007). Training Recurrent Neurocontrollers for Real-time Applications, IEEE Transactions on Neural Networks, 18(4), 1003-1015.
25. Schrauwen B., M. Wandermann, D. Verstraeten, J. J. Steil (2008). Improving Reservoirs Using Intrinsic Plasticity, Neurocomputing, 71, 1159-1171.
26. Si J., Y.-T. Wang (2001). On-line Learning Control by Association and Reinforcement, IEEE Transactions on Neural Networks, 12(2), 264-276.
27. Sutton R. S. (1988). Learning to Predict by Methods of Temporal Differences, Machine Learning, 3, 9-44.
28. Tohyama M., T. Patarinska, Z. Qiang, K. Shimizu (2002). Modeling of the Mixed Culture and Periodic Control for PHB Production, Biochemical Engineering Journal, 10, 157-173.
29. Werbos P. J. (1990). Backpropagation through Time: What It Does and How To Do It, Proceeding of the IEEE, 78(10), 1550-1560.

**Assoc. Prof. Petia Koprinkova-Hristova, Ph.D.**
E-mail: pkoprinkova@bas.bg

Petia Koprinkova-Hristova received M.Sc. degree in Biotechnics from the Technical University – Sofia in 1989 and Ph.D. degree on Process Automation from Bulgarian Academy of Sciences in 2001. Since 2003 she is an Associate Professor in the Institute of Control and System Research and from January 2012 – in the Institute of Information and Communication Technologies, Bulgarian Academy of Sciences. Her main research interests are in the field of intelligent control systems using mainly fuzzy, neuro-fuzzy and neural network approaches. Currently she is a member of European Neural Network Society (ENNS) executive committee for 2011-2016 and a member of the Union of Automatics and Informatics in Bulgaria.

**Assoc. Prof. Georgi Kostov, Ph.D.**
E-mail: george_kostov2@abv.bg

Georgi Kostov is an Associated Professor at the department "Technology of wine and brewing" at University of Food Technologies, Plovdiv. He received his M.Sc. in "Mechanical engineering" in 2007 and Ph.D. on "Mechanical engineering in food and flavor industry (Technological equipment in biotechnology industry)" in 2007 from University of Food Technologies, Plovdiv. His research interests are in the area of bioreactors construction, biotechnology, microbial populations investigation and modeling, hydrodynamics and mass transfer problems, fermentation kinetics.

**Assoc. Prof. Silviya Popova, Ph.D.**
E-mail: popova_silvia2000@yahoo.com

Silviya Popova received her M.Sc. in mathematics from University of Sofia, Bulgaria (1977) and Ph.D. on "New methods for automatization of videomicroscopy microbiological investigation" from the Bulgarian Academy of Sciences (2001). She habilitated as an Associate Professor in "Application of the Principles and Methods of Cybernetics in Different Areas of Science" in 2003. Currently she is with the Institute of Control and System Research, Department of Adaptive and Robust Control. She has won scholarships for specializations at University of Applied Science, Zittau, Germany (2002), Albert Ludvigs University, Freiburg, Germany (2006), University of Uppsala, Sweden (2006), University of Karlsruhe, Germany (2007). She is one of the founding members of the Bulgarian Women in Mathematics Society, established in 2002. Her research interests are in modeling, identification, estimation and control of biotechnological processes, adaptive control, neural networks and image processing.