Bioprocess Biosyst Eng (2006) 29: 39–47
DOI 10.1007/s00449-006-0054-3

ORIGINAL PAPER

B. Zelić · N. Bolf · Đ. Vasić-Rački

Modeling of the pyruvate production with *Escherichia coli*: comparison of mechanistic and neural networks-based models

Received: 10 February 2006 / Accepted: 17 February 2006 / Published online: 10 March 2006

Abstract Three different models: the unstructured mechanistic black-box model, the input–output neural network-based model and the externally recurrent neural network model were used to describe the pyruvate production process from glucose and acetate using the genetically modified *Escherichia coli* YYC202 *ldhA::Kan* strain. The experimental data were used from the recently described batch and fed-batch experiments [Zelić B, Study of the process development for *Escherichia coli*-based pyruvate production. PhD Thesis, University of Zagreb, Faculty of Chemical Engineering and Technology, Zagreb, Croatia, July 2003. (In English); Zelić et al. Bioproc Biosyst Eng 26:249–258 (2004); Zelić et al. Eng Life Sci 3:299–305 (2003); Zelić et al Biotechnol Bioeng 85:638–646 (2004)]. The neural networks were built out of the experimental data obtained in the fed-batch pyruvate production experiments with the constant glucose feed rate. The model validation was performed using the experimental results obtained from the batch and fed-batch pyruvate production experiments with the constant acetate feed rate. Dynamics of the substrate and product concentration changes was estimated using two neural network-based models for biomass and pyruvate. It was shown that neural networks could be used for the modeling of complex microbial fermentation processes, even in conditions in which mechanistic unstructured models cannot be applied.

Keywords Pyruvate · *Escherichia coli* · Unstructured "black-box" model · Input–output neural network-based model · Externally recurrent neural network model

B. Zelić (✉) · N. Bolf · Đ. Vasić-Rački
Faculty of Chemical Engineering and Technology,
University of Zagreb, Marulićev trg 19,
HR-10000, Zagreb, Croatia
E-mail: bzelic@fkit.hr
Tel.: +385-0-1-4597-146
Fax: +385-0-1-4597-133

Abbreviations c_A: acetate concentration (g L^{-1}) · $c_{A,0}$: acetate concentration in the feed (g L^{-1}) · c_G: gucose concentration (g L^{-1}) · $c_{G,0}$: glucose concentration in the feed (g L^{-1}) · c_P: pyruvate concentration (g L^{-1}) · $c_{P,MAX}$: critical pyruvate concentration above which reaction cannot proceed (g L^{-1}) · c_X: biomass concentration (g L^{-1}) · K_P: inhibition constant of Jerusalimsky (g L^{-1}) · K_S^A: monod growth constant for acetate (g L^{-1}) · K_S^G: monod growth constant for glucose (g L^{-1}) · m_A: Maintenance coefficient for growth on acetate (g g^{-1} h^{-1}) · m_G: maintenance coefficient for growth on glucose (g g^{-1} h^{-1}) · q_V: volumetric flow rate (L h^{-1}) · q_{VA}: volumetric flow rate of acetate (L h^{-1}) · q_{VG}: volumetric flow rate of glucose (L h^{-1}) · r_A: specific rate of acetate consumption (g g^{-1} h^{-1}) · r_G: specific rate of glucose consumption (g g^{-1} h^{-1}) · r_P: specific rate of pyruvate production (g g^{-1} h^{-1}) · t: time (h) · V: reaction (broth) volume (L) · $Y_{P/G}$: yield coefficient pyruvate from glucose (g g^{-1}) · $Y_{X/A}$: yield coefficient biomass from acetate (g g^{-1}) · $Y_{X/A,MAX}$: maximum yield coefficient biomass from acetate (g g^{-1}) · $Y_{X/G}$: yield coefficient biomass from glucose (g g^{-1}) · $Y_{X/G,MAX}$: maximum yield coefficient biomass from glucose (g g^{-1}) · u: input variables (–) · x: current output state (–) · μ: specific growth rate (h^{-1}) · μ_{MAX}: maximum specific growth rate (h^{-1})

Introduction

The quantitative comprehension of the dominant metabolic processes in the production strain, for instance substrate consumption, biomass growth, oxygen demand etc., is an important step in the development, optimization and scale-up of microbial fermentation processes. It is difficult to model and control bioprocesses owing to uncertainties and non-linear dynamics of biological phenomena. Therefore, black-box unstructured models were used due to their extreme simplicity

and applicability to a variety of modeling tasks. On the other hand, unstructured models required some rigorous constraints such as the balanced growth and one component cell system. Time-consuming and difficult experiments essential for an accurate estimation of the model parameter and cellular metabolism, which comprises many complex and interactive reactions irrespective of the unstructured models, would not produce unstructured models suitable for the estimator and control design [1, 2].

Mechanistic models such as the unstructured models of microbial fermentation processes based on mass and energy balances may be of a very high order. In the case of distributed parameter systems, the conversion of mechanistic model described by partial differential equations into a system of ordinary differential equations leads to a very high order model unsuitable for a direct application in an estimator framework. Therefore, it is important to create models that would be convenient for the estimator and the process design. Furthermore, the indicators of the bioprocess behavior, such as biomass and product concentration, are usually measured off-line, resulting in a very difficult recognition of undesirable fermentation, and leading to a considerable waste of time and resources. This problem has led to the development of a range of the so-called ''software sensors'', which utilize mathematical models and algorithms, together with the on-line available measured process variables to estimate the behavior of the fermentation process [3].

One of the main obstacles to the effective bioprocess operation is a lack of on-line information regarding the bioprocess condition. They are mostly model based and as such require a good description of the mathematical process. The excellent neural network characteristics such as the treatment of non-linear systems, noisy and approximate data, learning from the past data in order to adapt to a changing environment and a prompt execution of actions once the network has been trained offer a great potential in biosensor data processing, variable prediction, optimization and advance control of dynamic bioprocess [4–9]. Inevitably, there must be some simplifications in the model structure, leading to a simplified model that does not truly represent the real complexity and non-linearity of the process. Nevertheless, the accuracy provided by artificial neural networks turns them into an attractive approach for the consideration of bioprocess variable estimation [10]. The variety of available neural network architectures allows us to deal with a wide range of bioprocess modeling and control problems. In comparison to other empirical models, neural networks are relatively less sensitive to noise and incomplete information, thus facilitating the management of a higher level of uncertainty when applied in process control problems [11].

The advantage of using a neural network to simulate a process lies in the fact that it represents a quick and reliable way of dynamic performance prediction. It can also be continuously updated. Two groups of methods can be found in references that can be used for a dynamic process modeling by means of a neural network. The first group is the so-called input–output models [12], and the second group is recurrent neural networks [13].

The advantage of using the first group for the approximation of the dynamic process is the possibility of using any static neural net, since the dynamic behavior of the process is enforced by the use of history data. The problem with these models is a need to detrend the measurement data in order to achieve the feasibility to model long-term dependencies. Hence, the second group of methods is often used. This group includes the so-called recurrent neural nets [13]. Recurrent nets are usually internally recurrent structures such as the Elman network [14], the Jordan network [15] or the classical dynamical system described by Perreto and Niez [16].

Dynamical neural models can be internally or externally recurrent. Internally recurrent neural models are differential equation systems requiring parameter determination. There are a number of methods for network training in references. These methods are relatively complex because they have to solve a series of practical problems such as the possible numerical instability or the information loss problem in terms of distant time characteristics (gradient vanishing). Consequently, Nerrand developed an externally recurrent neural network [17].

Neural networks were shown to be superior to mechanistic models in the description of bioprocesses in many lab-scale and industrial, real or simulated processes. The application of neural network ranges from the estimation of biomass and product (penicillin, amino acids, secondary metabolites) concentrations [18–20] to on-line control and optimization of fed-batch fermentation and enzyme production processes [21, 22].

The purpose of this paper is to compare the experimental results of batch, fed-batch and continuous experiments with the simulation results of mechanistic unstructured ''black-box'' model and the simulation results of neural network-based models in order to test the effectiveness and accuracy of neural networks in bioprocess modeling and estimation. The pyruvate production process from glucose and acetate, using genetically modified *Escherichia coli* YYC202 *ldhA::Kan* strain, was used as a model system. This strain is completely blocked in its ability to convert pyruvate into acetyl–CoA or acetate [23, 24]. The production strain, mediums, process conditions and cultivation methods, as well as the experimental results, are described elsewhere [1, 2, 23, 24].

Model description

In order to simplify the model, the bioconversion of glucose to pyruvate is regarded as a one-step-enzymatic reaction:

$$C_6H_{12}O_6 \xrightarrow{\textit{Escherichia coli}, \text{ acetate}, O_2} 2C_3H_4O_3 + 2H_2O$$

The following constraints were used to define all the models: glucose and acetate are only limiting substrates; cell growth occurs only in the presence of both substrates; there was no oxygen effect on biomass growth and pyruvate production; product formation kinetics should combine growth-associated and non-growth-associated characteristics; bioconversion of glucose to pyruvate was assumed to be a one-step-enzymatic reaction; both biomass growth and pyruvate production were inhibited by high pyruvate concentrations; the viscosity of the reaction mixture remains constant during experiments; and the potential mixing effects of the highly concentrated feed with the cultivation medium are neglected in order to protect the model simplicity.

Mechanistic unstructured model

Modeling of the pyruvate production process using genetically modified *E. coli* YYC202 *ldhA::Kan* strain was already presented [1, 2]. The model equations for the batch and fed-batch pyruvate production process, which describe biomass growth, pyruvate formation, glucose and acetate uptake, followed by the volume change in the case of fed-batch process, are represented by the following set of differential equations (Eqs. 1–5).

$$\frac{dc_X}{dt} = -\frac{q_V}{V} \times c_X + \mu \times c_X \quad (1)$$

$$\frac{dc_G}{dt} = -\frac{q_V}{V} \times c_G + \frac{q_{VG}}{V} \times c_{G,0} - r_G \times c_X - r_P \times c_X \quad (2)$$

$$\frac{dc_A}{dt} = -\frac{q_V}{V} \times c_A + \frac{q_{VA}}{V} \times c_{A,0} - r_A \times c_X \quad (3)$$

$$\frac{dc_P}{dt} = -\frac{q_V}{V} \times c_P + r_P \times c_X \times Y_{P/G} \quad (4)$$

$$\frac{dV}{dt} = q_{VG} + q_{VA} = q_V \quad (5)$$

where, c_X, c_G, c_A and c_P are the biomass, glucose, acetate and pyruvate concentrations, respectively, V is the biosuspension volume, q_V is the time-dependent overall volumetric flow rate, q_{VG} and q_{VA} are glucose (G) and acetate (A) analogues, and $c_{G,0}$ and $c_{A,0}$ are glucose and acetate concentration in the feed. The kinetics of biomass growth (μ), pyruvate formation (r_P), glucose uptake (r_G), acetate uptake (r_A) are described as follows (Eqs. 6–9):

$$\mu = \mu_{MAX} \times \frac{c_G}{K_S^G + c_G} \times \frac{c_A}{K_S^A + c_A} \times \frac{K_P}{(c_P + K_P)} \quad (6)$$

$$r_P = \left(\alpha \times \frac{dc_X}{dt} + \beta \times c_X\right) \times \left(1 - \frac{c_P}{c_{P,MAX}}\right) \quad (7)$$

$$r_G = \frac{\mu}{Y_{X/G}} \quad (8)$$

$$r_A = \frac{\mu}{Y_{X/A}} \quad (9)$$

The model combines the growth inhibition by pyruvate described by Jerusalimsky approach [25], and the pyruvate inhibited product formation described by a modified Luedeking-Piret/Levenspiel term [2]. The model developed by Jerusalimsky represents an approximate analogy to the non-competitive substrate inhibition, which is often used in pure enzyme kinetic models. The yields such as biomass/glucose, $Y_{X/G}$ and biomass/acetate, $Y_{X/A}$ were assumed to be the functions of biomass growth and maintenance energy demand (Eqs. 10–11).

$$Y_{X/G} = \frac{Y_{X/G,MAX} \times \mu}{Y_{X/G,MAX} \times m_G + \mu} \quad (10)$$

$$Y_{X/A} = \frac{Y_{X/A,MAX} \times \mu}{Y_{X/A,MAX} \times m_A + \mu} \quad (11)$$

Neural network-based model of the pyruvate production process

The objective was to estimate the biomass and pyruvate concentrations that are difficult to measure on-line. The estimation algorithms were developed using the two above-mentioned techniques: the input–output neural network-based model and the recurrent neural network model. Both of the models were developed, fitted, simulated and tested using real pyruvate production experimental data. The standard input–output "black-box" model, shown in Fig. 1a, was previously described [26].

All data are divided into train, test and validation sets. The test set is indirectly used in the model construction, whereas the validation set is completely independent of the model construction. Twenty percent of all data are chosen randomly from the data for the validation set, which is specified by the software package. Although the explicit training of the model uses only the training set, the performance on the test set is used to

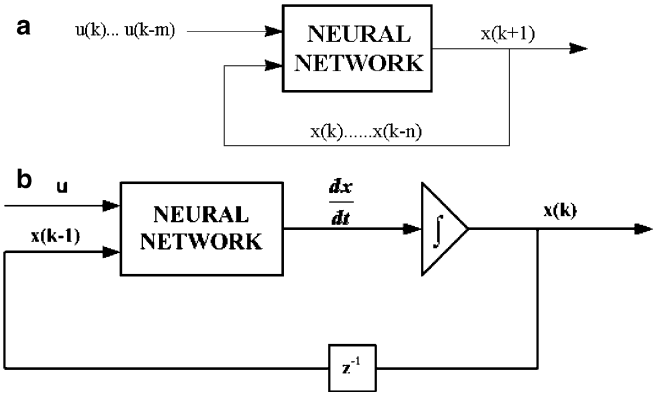

Fig. 1 **a** Input–output neural network. **b** Continuous externally recurrent neural network

guide choices during the model construction. Several models are built using the train–test set, and then the final model is chosen, based on its performance on the validation set. Yet another level of validation set for the final performance assessment is used with another fed-batch experiment.

Continuous externally recurrent neural network

To improve the possibility of the neural approximation and to prevent potential integration instability, the continuous version of the externally recurrent network is used (Eq. 12):

$$\frac{dx}{dt} = f(x,u) \tag{12}$$

In order to gather the data necessary for the neural network training, the right-hand side of the system (7) must be evaluated. This task can be accomplished by a numerical derivation of the original measurement data values with respect to time by one of the known techniques [27]. Therefore, it is simpler to use the externally recurrent dynamical neural model in the state space. Such a model is also called a neural model in the state space or a canonical neural model, with the mathematical description as follows (Eq. 13):

$$f_k(x_k,u_k) \approx \left(\frac{\Delta x}{\Delta t}\right)_k \tag{13}$$

The structure of the externally recurrent neural network is shown in the Fig. 1b. It can be seen that the neural network represents the right-hand side of the Eq. (12) as a function of state and input variables.

During the recurrent neural network construction procedure, several structures of feed forward networks were evaluated with different sizes of the time window. The size of the time window is optimized by the use of the trial and error method [28]. After the training and testing procedure, the chosen one has three equally spaced time delay units of 0.05 h for input variables.

Building of the neural network model by cascade learning

Cascade learning based on the cascade-correlation learning paradigm [29] is developed for the neural network construction. Cascade learning starts off with no hidden nodes. The only connection is a direct connection from the input layer (and bias) to the output layer. Hidden nodes are added one at a time, and the purpose of each new hidden node is to predict the current remaining output error in the network. Hidden nodes receive the input from all previous hidden nodes as well as from the input buffer; in other words, the hidden layer has cascaded connections.

In our application the cascade-correlation algorithm works as follows:

1. Train the direct connections from the input layer and bias to the output layer. Train until the RMS (root mean square) output error stabilizes.
2. Iterate on the following steps:
   - Train a new hidden node so as to maximize a measure of the correlation between its output and the residual error at the output for the current training vector. The untrained hidden node is referred to as the ''candidate''. When training has stabilized, or after a given number of training iterations, learning is permanently disabled for the incoming connections to that node. At this point, the hidden node is said to be ''tenured'';
   - Connect the newly tenured hidden node to all nodes in the output layer, and randomly initialize the weights at those connections;
   - Train all the weights at all connections from the input layer, bias and tenured hidden nodes to the output layer. Train until the RMS output error stabilizes.

The iterative steps are repeated until the performance of the network (RMS error measured on a test set) no longer shows any improvement.

Results and discussion

Model development and training of the neural networks

The development of the unstructured mechanistic model, estimation of parameters and validation of the model for the bioconversion of glucose to pyruvate using genetically modified *E. coli* strain were described elsewhere [2]. All experiments were carried out in a 7.5 L bioreactor at a temperature of 37°C. Sufficient aeration (dissolved oxygen[DO]≥40%) was obtained by vigorous stirring (200–1,800 rpm), airflow rate (1–10 L min^{-1}) and reactor overpressure (0.2–0.8 bar).

The estimation of model parameters and model validation were performed using the data from fed-batch fermentations carried out at the constant glucose feed rates (glucose concentration in the feed medium of 700 g dm^{-3}) of 10, 20 and 30 cm^3 h^{-1} . On the other hand, acetate was fed according to the previously developed feeding strategy [23]. In this heuristic approach, the acetate consumption was calculated based on the on-line estimated CO_2 production rate. The acetate concentration in the acetate feed medium was 109 g dm^{-3} . The model identification was realized by the least-square fit, and the model discrimination was based on the model selection criterion (MSC), respectively.

Using this modeling approach, the acceptable model prediction for pyruvate formation was achieved (Figs. 2, 3 and 4). Unfortunately, the dynamics of pyruvate formation is just one of the essential variables to model process alternatives and scale-up. In the case of biomass,

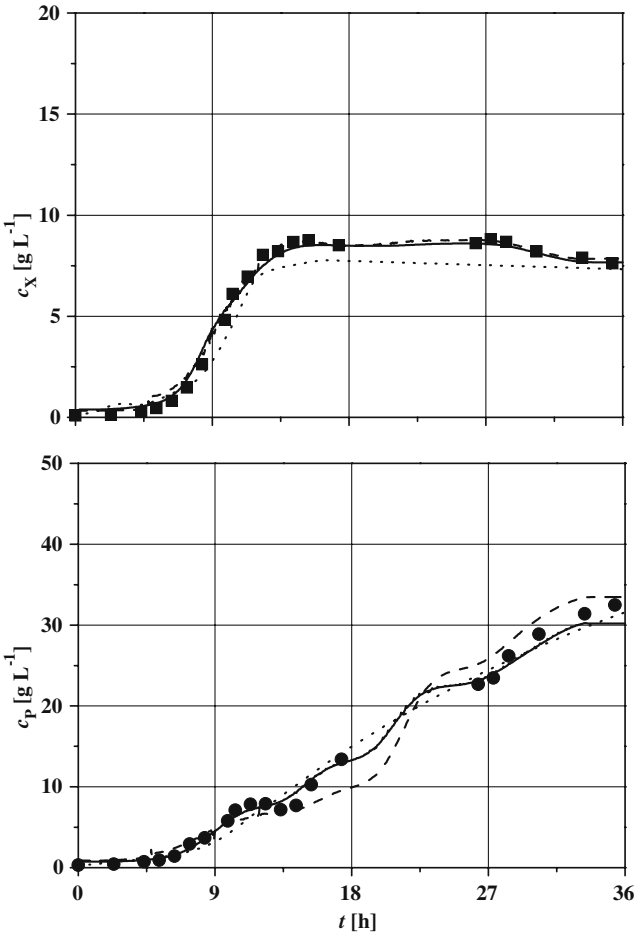

Fig. 2 Data obtained by model simulation (unstructured black-box model ·····; neural network model - - -; externally recurrent neural network model —) and experimental data for biomass (*filled square*) and pyruvate (*filled circle*) concentration in fed-batch process at constant glucose volumetric flow rate of $q_{VG} = 10$ cm^3 h^{-1}

acetate and glucose, further studies must be performed to increase the model predictive quality. The need to incorporate additional aspects of energy demand or by-product formation was identified, necessitating the extension of the extant model by the structured modeling terms. Nevertheless, the identified unstructured model was qualified as a promising tool for modeling studies as well as for further, more detailed kinetic modeling approaches. However, these kinds of models are very sensitive to noises and disturbances, e.g., process failures such as in the feed pump, level controller, valve or coolant flow. The complex interactions between kinetics and fluid mixing and the difficulty of maintaining strict sterility over a long period are additional points that cannot be covered by the use of the unstructured model.

Unstructured models are, in most of the cases, inapplicable to process monitoring and control of large-scale fermentations [30]. Therefore, neural network-based models, which have been shown to be very useful in describing bioreactor behavior [18, 20–22] in both real and simulated industrial conditions, were used to represent pyruvate production process. To train externally recurrent neural networks, it was necessary to generate additional data and derive original concentration trajectories. For that purpose the negative exponential smoothing method was used, and the obtained sets of data were applied as training sets. The smoothing method weighs the data contained in a window surrounding the smoothing location. The radius of this window, the so-called bandwidth radius, was constant and dependent on the data set time window (each data set was divided into 500 intervals). The negative exponential smoothing method applies a Gaussian weight function to weigh the data and the quadratic fit. The applied smoothing method parameters encompassed the sampling proportion 0.1 and the polynomial degree 1–4. The network was trained by the adaptive gradient learning algorithm.

The input layer consists of four state variables (biomass, glucose, acetate and pyruvate concentrations) and two input variables (glucose and acetate feed rates). Since the glucose and acetate flow rates were maintained constant, the volume was not introduced as a state variable in this modeling approach. In the general case where the flows vary with time volume, it should be included in the structure of neural networks. The output layer includes the prediction of future values for biomass (first network) and pyruvate concentration (second network). Neural networks were built afterward within hidden layers by cascade learning, adding hidden nodes as described in "Building of the neural network model by cascade learning" section. This procedure was undertaken using the NeuralWare software package *Predict* and *ProfII/Plus* running within Microsoft Excel [29]. The final network architectures applied for the prediction of biomass and pyruvate concentration have one hidden layer consisting of four and eight hidden nodes, respectively.

The number of samples used in the training set (680) and the specified number of hidden nodes (4 and 8) satisfied the criteria that the number of freedom is within a half or a quarter of the number of samples used in the training procedure [12, 31].

Validation of the models

Being aware that kinetic modeling is a difficult task for bioprocesses, a succession of decisions must be made in respect of those systematic methodologies that are still lacking. Therefore, it is important to validate a model by using data other than those applied to identify parameters. Furthermore, since the development of an appropriate feeding strategy is critical in fed-batch fermentation, the unstructured model and both neural network- based models, respectively, were validated on experimental data set collected in the fed-batch fermentation with constant acetate flow rate $q_{VA} = 1.3$ cm^3 h^{-1}, but glucose was not fed in the

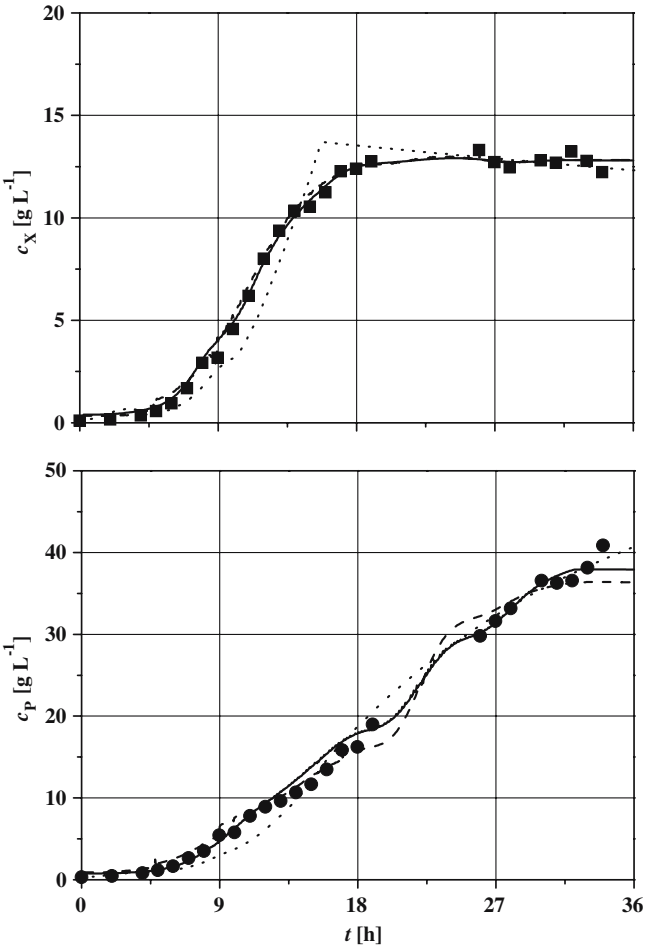

Fig. 3 Data obtained by model simulation (unstructured black-box model ·····; neural network model - - -; externally recurrent neural network model —) and experimental data for biomass (*filled square*) and pyruvate (*filled circle*) concentration in fed-batch process at constant glucose volumetric flow rate of $q_{VG} = 20$ cm^3 h^{-1}

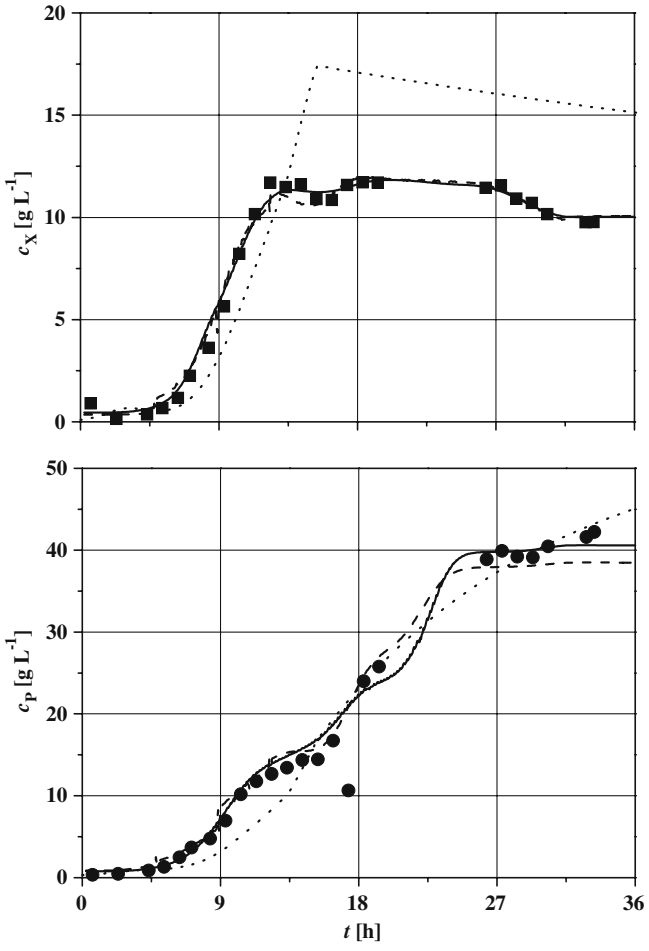

Fig. 4 Data obtained by model simulation (unstructured black-box model ·····; neural network model - - -; externally recurrent neural network model —) and experimental data for biomass (*filled square*) and pyruvate (*filled circle*) concentration in fed-batch process at constant glucose volumetric flow rate of $q_{VG} = 30$ cm^3 h^{-1}

bioreactor. In this experiment, the initial glucose concentration was $c_G = 23$ g dm^{-3}, and glucose was in excess during the experiment (data not shown).

Obviously, the unstructured black-box model is unable to predict process dynamics quantitatively, e.g., it is unable to predict biomass and pyruvate concentrations with acceptable accuracy (Fig. 5). Interestingly, pyruvate levels predicted by those models were approximately 50% lower, and biomass levels were 50% higher than the experimental observations. One of the main reasons for this kind of behavior of the unstructured model simulation results is the impossibility to use this kind of model for the description of experimental results collected in the conditions different from those used for the estimation of model parameters. Namely, the initial glucose concentration in the fed-batch experiment with constant glucose feed rate ($c_{G,0} = 10$ g dm^{-3}) was twofold lower than in the fed-batch experiment with constant acetate feed rate ($c_{G,0} = 20$ g dm^{-3}). Furthermore, during the fed-batch phase of the experiment performed with constant glucose feed rate, the glucose concentration level was under the detection limit. In the fed-batch experiment with constant acetate feed rate, the glucose was in excess during complete fermentation, which obviously resulted in a different response of complex and various intracellular reactions, which were not covered with the present unstructured "black-box" model. This statement appears more reasonable after the comparison of experimental and simulation results for the batch experiment. The initial glucose and acetate concentration in the batch experiment were 17.4 g dm^{-3} and 0.5 g dm^{-3}, respectively. In the batch experiment, the differences observed in the fed-batch experiment with the constant acetate feed are more intensive for both the biomass and pyruvate concentration profiles investigated (Fig. 6).

On the other hand, the applied neural network-based models are able to follow the dynamic behavior of the process quite well. Figures 2–4 show the results achieved with the data set used for network training. After the

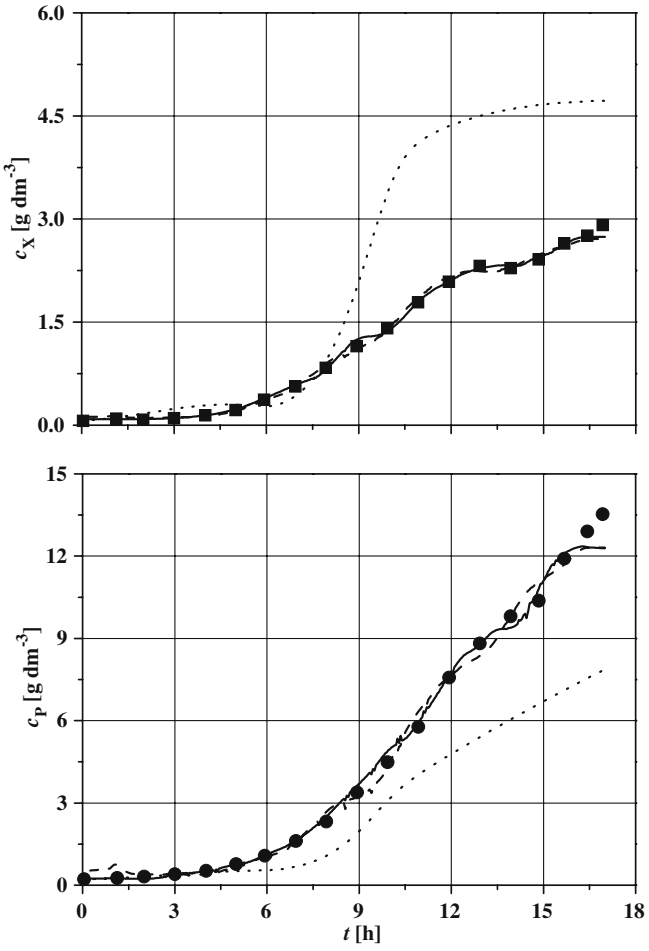

Fig. 5 Data obtained by model simulation (unstructured black-box model ·····; neural network model - - -; externally recurrent neural network model —) and experimental data for biomass (*filled square*) and pyruvate (*filled circle*) concentration in fed-batch process at constant acetate volumetric flow rate of $q_{VA} = 1.3$ cm^3 h^{-1}

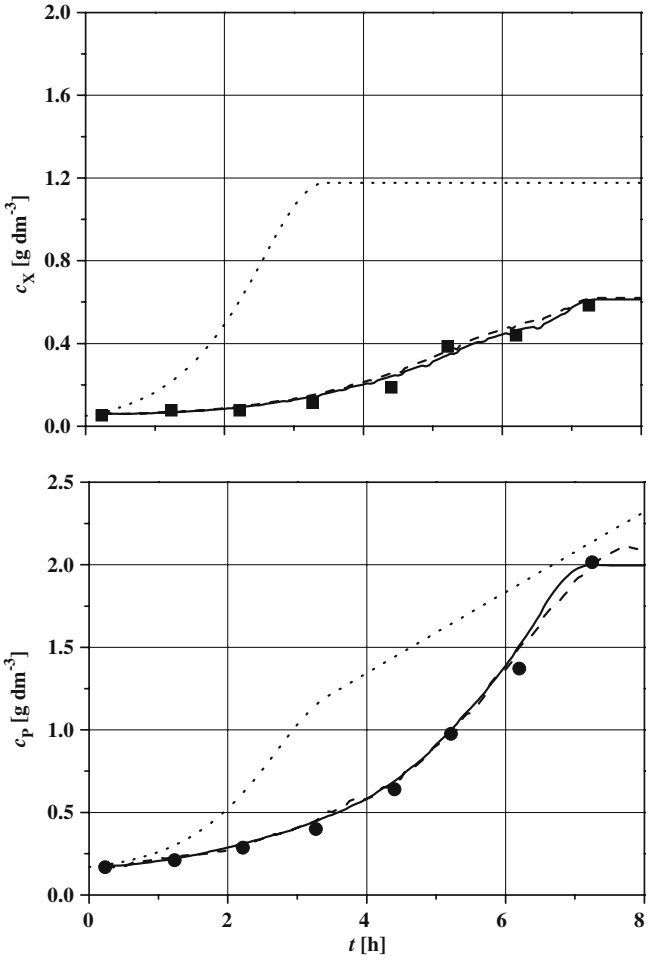

Fig. 6 Data obtained by model simulation (unstructured black-box model ·····; neural network model - - -; externally recurrent neural network model —) and experimental data for biomass (*filled square*) and pyruvate (*filled circle*) concentration in batch process

examination of the data (glucose feed rate 10, 20 and 30 cm^3 h^{-1}), it can be observed that the biomass model follows the experimental data fairly accurately. As for the pyruvate concentration, the input/output neural network model slightly oscillates, which can be linked to an imbalance in the accessibility of the experimental data. Moreover, a certain deviation in the pyruvate concentration can be seen even prior to the end of the experiment for the input–output neural model.

When applying the neural network models to the constant acetate flow experiment (Fig. 5), a slight oscillation in the models and deviations can still be seen by the end of the experiment. Nevertheless, the experimental data are gathered in a more balanced manner, so the model follows the given values fairly well. For the batch experiment, both the models follow the experimental data accurately (Fig. 6), although there are less of them but with shorter experimentation time.

Table 1 shows the comparison of average absolute errors and maximum absolute errors for neural network models. In this context, the training set is a set of points that are used to fit the model parameters. The test set is used as a part of the model-building process to prevent

Table 1 Parameters of input–output (NN1) and externally recurrent neural network (NN2) performances. Avg. Abs., the average absolute error between the target output and the prediction; Max. Abs., maximal absolute error between the target output and the prediction

	Data Set	Avg. Abs.	Max. Abs.
$c_{X(NN1)}$	Train	0.8346	4.8319
	Test	0.8156	4.6155
	Validation	0.8543	4.9350
$c_{P(NN1)}$	Train	0.6072	1.8212
	Test	0.5983	1.8620
	Validation	0.6045	1.8740
$c_{X(NN2)}$	Train	0.8127	4.0158
	Test	0.8125	4.0167
	Validation	0.8132	4.0137
$c_{P(NN2)}$	Train	0.1210	2.3261
	Test	0.1240	0.6839
	Validation	0.1236	1.2294

Table 2 Residual sum of squares for unstructured mechanistics model (UMM), input–output (NN1) and externally recurrent neural network (NN2)

Experiment	Residual sum of squares		
	UMM	NN1	NN2
Batch	2.05×10^{3}	35.79	35.51
Fed-batch (constant acetate flow rate $q_{VA} = 1.3$ cm^3 h^{-1})	8.37×10^{3}	64.65	50.90
Fed-batch (constant glucose flow rate $q_{VG} = 10$ cm^3 h^{-1})	$1.99\ 10^{2}$	24.75	67.38
Fed-batch (constant glucose flow rate $q_{VG} = 20$ cm^3 h^{-1})	3.14×10^{2}	44.23	67.56
Fed-batch (constant glucose flow rate $q_{VG} = 30$ cm^3 h^{-1})	1.02×10^{3}	74.14	101.55

over-fitting. The validation set is used as an additional independent validation test set. According to these measures, the externally recurrent network shows better performance. Table 2 shows the sum of the squared error comparison for the three applied models. The table clearly shows the applicability of the neural network model in comparison to the unstructured mechanistical model. Both networks have proven to be equally good in the batch process, while the externally recurrent network has been better in validating the fed-batch. The input/output neural network has shown slightly better performance on the training set, but the externally recurrent neural network model is also applicable.

Based on our experience, the externally recurrent neural network should have some advantage over input–output network. This advantage is due to the fact that the integration can be performed with an arbitrary method for the integration of differential equation systems. In practice, that means capability to deal with stiff and instable systems. The obvious disadvantage of the externally recurrent network is a necessity to do a considerable amount of data manipulation in order to prepare data for network learning, such as smoothing and derivation. The method is useful only for an autonomous (time invariant) process where time is not contained explicitly in the process model, so any static neural net can be used.

Conclusion

The unstructured mechanistic ''black-box'' model for the bioconversion of glucose to pyruvate using *E. coli* YYC202 *ldhA::Kan* strain cannot fit adequately the experimental data obtained in the different processes. The developed model is able to give some basic information about the process and possible process optimization, and scale-up further model enlargement on structural modeling terms.

The developed neural network-based models are relatively simple methods for real-world application. The results obtained using the neural networks show satisfactory compliance with experimental values. The performances of the neural networks show conclusively that they have the potential to be implemented as an online state-estimator, facilitating the control of pyruvate production, and also be used in process optimization.

Acknowledgments All experiments were performed at the Institute of Biotechnology 2, Research Center Jülich, Germany. The authors are indebted to Prof. Dr Christian Wandrey and Dr Ralf Takors for generous support of this work. This work was partially funded by the Croatian Ministry of Science, Education and Sport, contract grant number 0125 021.

References

1. Zelić B (2003) Study of the process development for *Escherichia coli* based pyruvate production. PhD Thesis, University of Zagreb, Faculty of Chemical Engineering and Technology, Zagreb, Croatia (In English)
2. Zelić B, Vasić-Rački Đ, Wandrey C, Takors R (2004) Modeling of the pyruvate production with *Escherichia coli* in a fed-batch bioreactor. Bioproc Biosyst Eng 26:249–258
3. Prasad V, Bequette BW (2003) Nonlinear system identification and model reduction using artificial neural networks. Comp Chem Eng 27:1741–1754
4. DiMassimo C, Montague GA, Willis MJ, Morris AJ, Tham MT (1992) Enhancing industrial bioprocess monitoring through artificial neural networks. IFAC Modeling and Control of Biotech Proc, Keystone, Colorado, 395–398
5. Linko P, Zhu YH (1991) Neural network programming in bioprocess variable estimation and state prediction. J Biotechnol 21:253–270
6. James S, Legge R, Budmann H (2002) Comparative study of black-box and hybrid estimation methods in fed-batch fermentation. J Proc Control 12:113–121
7. Linko S, Luopa J, Zhu YH (1997) Neural networks as ''software sensors'' in enzyme production. J Biotechnol 52:257–266
8. Martin GD (1997) Consider soft sensors. Chem Eng Prog 66–70
9. Dochain D, Perrier M (1997) Dynamical modeling, analysis, monitoring and control design for nonlinear bioprocesses. Adv Biochem Eng 56:149–197
10. Kapil GG, Mehra S, Gomes J (2005) On-line adaptation of neural network for bioprocess control. Comp Chem Eng 29:1047–1057
11. Baughman DR, Liu YA (1995) Neural networks in bioprocessing and chemical engineering. Academic, San Diego, CA
12. Masters T (1995) Neural, novel & hybrid algorithms for continually running fully recurrent neural networks. Wiley, New York
13. Gosak D, Vampola M (2001) Generic neural model control of batch distillation process. IchemE symposium ''Advances in Process Control'', CD-ROM
14. Elman JL (1990) Finding structure in time. Cogn Sci 14:179–211
15. Jordan MI (1986) Attractor dynamics and parallelism in a connectionist sequential machine. The eight annual conference of the cognitive science society, Amherst, USA, 531–546
16. Perretto P, Niez JJ (1986) Stochastic dynamics of neural networks. IEEE Trans Syst Man Cybern 16:73–83
17. Nerrand O (1994) Training recurrent neural networks: why and how? An illustration in dynamical process modelling. IEEE Trans Neural Netw 5:178–184

18. Warnes MR, Glassey J, Montague GA, Kara B (1998) Application of radial basis function and feedforward artificial neural networks to the *Escherichia coli* fermentation process. Neurocomputing 20:67–82
19. Bachinger TH, Martensson P, Mandenius CF (1998) Estimation of biomass and specific growth rate in a recombinant *Escherichia coli* batch cultivation process using a chemical multisensor array. J Biotechnol 60:55–66
20. Hajmeer MN, Basheer IA (2003) A hybrid Bayesian–neural network approach for probabilistic modeling of bacterial growth/no-growth interface. Int J Food Microbiol 82:233–243
21. Lee J, Lee SY, Park S, Middelberg APJ (1999) Control of fed-batch fermentations. Biotechnol Adv 17:29–48
22. Vlassides S, Ferrier JG, Block DE (2001) Using historical data for bioprocess optimization: modeling wine characteristics using artificial neural networks and archived process information. Biotechnol Bioeng 73:55–68
23. Zelić B, Gerharz T, Bott M, Vasić-Rački Đ, Wandrey C, Takors R (2003) Fed-batch process for pyruvate production by recombinant *Escherichia coli* YYC202 strain. Eng Life Sci 3:299–305
24. Zelić B, Gostović S, Vuorilehto K, Vasić-Rački Đ, Takors R (2004) Process strategies to enhance pyruvate production with recombinant *Escherichia coli*: from repetitive fed-batch to in situ product recovery with fully integrated electrodialysis. Biotechnol Bioeng 85:638–646
25. Roels JA (1983) Energetics and kinetics in biotechnology. Elsevier, Amsterdam
26. Sander A, Skansi D, Bolf N (2003) Heat and mass transfer models in convection drying of clay slabs, Ceramic Int 29:641–653
27. De Boor C (1978) A practical guide to splines. Springer Verlag, Berlin Heidelberg New York
28. Novakovic B, Majetic D, Siroki M (1998) Artifical neural networks. University Press, Zagreb
29. Neuralware, Neural computing—A technology handbook for professional II/plus and neuralworks explorer, NeuralWare Inc., Pittsburgh, USA, 1996
30. Patnaik PR (1997) A recurrent neural network for a fed-batch fermentation with recombinant *Escherichia coli* subject to inflow disturbances. Proc Biochem 32:391–400
31. Bishop CM (1995) Neural network for pattern recognition. Oxford University Press, New York