Modelling and control of nonlinear system

David Sámek, Petr Dostál

Abstract
This paper is focused on modelling and control of nonlinear system of two interconnected cylindrical tanks. The first part of paper is devoted to brief introduction to theoretical background. Then, the mathematical model of the real-time laboratory system is derived. The experimental measurements were made after the mathematical modelling in order to obtain parameters of the mathematical model. After that, the simulation model in the Matlab/Simulink was developed using S-Function feature. The presented controller is based on model predictive control method using artificial neural network as a predictor. In this contribution are compared two different artificial neural networks.

Keywords: model predictive control, artificial neural networks, nonlinear system

Introduction

The increasing demand on production systems and the need for constant changes in the normal operating conditions of industrial processes is reflected in the development of more reliable control systems. In fact, new production systems represent complex and highly nonlinear engineering problems where the conventional linear control methods can hardly be successfully applied. Additionally, the overall complexity of the control problem is increased when considering physical, environmental or safety constraints to be followed during process operation. Therefore, the urge in finding faster and more reliable solutions for highly nonlinear control problems in constrained environments is one of today's trends in modern control.

Model predictive control (MPC) [1] is a very attractive concept for the development and tuning of nonlinear controllers in the presence of input, output or state constraint. The MPC controllers contain predictors that predict the controlled system output. One of successful approaches for prediction is usage of an artificial neural network (ANN).

Many predictive control techniques based on MPC that use artificial neural network as a predictor are established on multilayer feed-forward neural networks [2], [3]. In spite the multilayer feed-forward neural networks (MFFNNs) have many advantages such as simple design and scalability they have also many drawbacks such as long training times and choice of an appropriate learning stop time (the over-learning versus the early stopping).

Nevertheless, there are quite a number of types ANNs suitable for the modelling and prediction. Recurrent artificial networks are very promising for modelling and prediction of nonlinear systems [4], [5], [6]. Radial Basis Function (RBF) artificial neural networks are popular for their rapid training [7], [8].

In this article adaptive linear networks (ADALINEs) are presented. ADALINEs offer even faster training and shorter computational times [9]. As a comparative method the MFFNN was chosen.

This paper is organized in the following way. Section 1 presents brief introduction to Model Predictive Control methods. Section 2 deals with the description of Adaptive linear networks. Section 3 presents short introduction to Multilayer feed-forward neural networks. The mathematical description of the controlled system is provided in Section 4. In the Section 5 we explain the structure design of the predictors that were used in the paper. This section is followed by simulations and results in Section 6. The paper is concluded by results discussion in Section 7 and with some final remarks in conclusion.

1. Model predictive control using ANN

There are various approaches to predictive control by artificial neural networks. Generally we can say that these methods use ANN as the plant model in order to get its output predictions. The most used approach is model predictive control [1]. MPC is a broad control strategy applicable to both linear and nonlinear processes.

The main idea of MPC algorithms is to use a dynamical model of process to predict the effect of future control actions on the output of the process. Hence, the controller calculates the control input that will optimize the performance criterion over a specified future time horizon:

$$J = \lambda \sum_{j=N_1}^{N_2} \left[y_r(k+j) - \hat{y}(k+j) \right]^2 + \\ + \rho \sum_{j=1}^{N_u} \left[u_t(k+j-1) - u_t(k+j-2) \right]^2 \tag{1}$$

where N_1, N_2 and N_u define horizons over which the tracking error and the control increments are evaluated. The u_t variable is the tentative control signal, y_r is the desired response and $\hat{y}$ is the network model response.

The parameters λ and γ determine the contribution that the sum of the squares of the future control errors and the sum of the squares of the control increments has on the performance index.

Typically the receding horizon principle is implemented, which means that after the computation of optimal control sequence, only the first control action is implemented. Then the horizon is shifted forward one sampling instant and the optimization is again restarted with new information from measurements. This methodology is adopted in this paper.

In cases where the model of the process is given as a nonlinear combination of the process inputs (e.g. ANN), the solution of the standard constrained MPC is necessarily more complex. A schematic configuration of such control structure is presented in figure 1. Due to nonlinear nature of ANN prediction model a nonlinear optimization problem must be solved through a numerical algorithm.

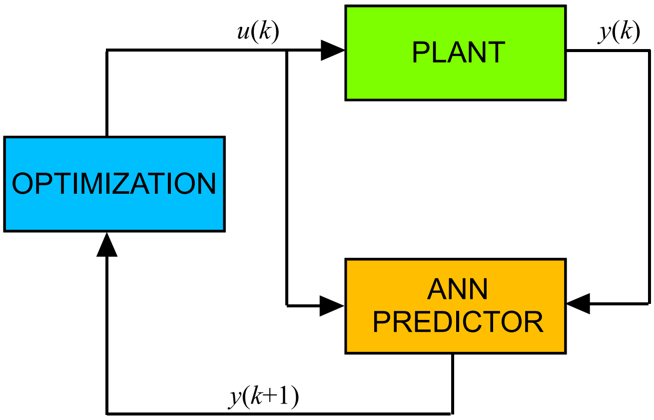

Fig.1 Principle of MPC using ANN

2. Adaptive linear networks

ADALINE was described by B. Widrow and M.E. Hoff as an adaptive threshold logic element in 1960 [9]. Though, the original version of ADALINE had only simple two-state threshold transfer function with the range of function {-1;+1}, nowadays ADALINE is also used with linear transfer function [10]. Although this structure has limited skills, it is possible to connect more of ADALINEs together to obtain MADALINE (Multiple ADALINE).

Despite the fact that ADALINEs are able to solve only linearly separable problems, it has been shown in practice that they can approximate nonlinear functions with the sufficient accuracy while using enough number of neurons [11]. Because of their main advantage that is the very fast learning, they have many practical applications, e.g. noise reduction, signal processing and signal prediction in control and communication systems. The learning procedure is based on an iterative search process, where performance feedback is used to guide the search process. In other words, a designer "trains" the system by "showing" it examples of inputs and the respective outputs. In this way, the system competence is directly and quantitatively related to the amount of experience the system was given.

The most popular learning method is simple LMS (Least Mean Square) algorithm [9], often called the Widrow-Hoff Delta Rule, which is adopted in this paper. This method is based on the minimization of Mean Square Error (MSE) which is for j-th ADALINE defined:

$$E_j = \frac{1}{n}\sum_{i=1}^{n}\left(\tau(i) - y_{out}(i)\right)^2 \tag{2}$$

where $\tau(i)$ is i-th target neuron output, $y_{out}(i)$ is output from i-th ADALINE, n is number of training data. For Multiple ADALINE (MADALINE) with m neurons is necessary to compute global MSE:

$$E = \frac{1}{m}\sum_{j=1}^{m} E_j \tag{3}$$

New weighting matrix $\boldsymbol{W}(k+1)$ and bias vector $\boldsymbol{b}(k+1)$ in the step $k+1$ equal to (for ADALINE $\boldsymbol{W}$ has only one row and $\boldsymbol{b}$ is scalar):

$$\boldsymbol{W}(k+1) = \boldsymbol{W}(k) + \alpha \cdot \boldsymbol{\varepsilon}(k) \cdot \boldsymbol{u}_{in}(k) \tag{4}$$

$$\boldsymbol{b}(k+1) = \boldsymbol{b}(k) + \alpha \cdot \boldsymbol{\varepsilon}(k) \tag{5}$$

where $\boldsymbol{W}(k)$ is previous weighting matrix, $\boldsymbol{b}(k)$ is previous bias vector, α is learning rate from the interval <0; 1>, $\boldsymbol{u}_{in}(k)$ is the vector of input data and $\boldsymbol{\varepsilon}(k)$ is output error in the step k:

$$\boldsymbol{\varepsilon}(k) = \boldsymbol{\tau}(k) - \boldsymbol{y}_{out}(k) \tag{6}$$

where $\boldsymbol{y}_{out}(k)$ is actual output from MADALINE:

$$\boldsymbol{y}_{out}(k) = \boldsymbol{b}(k) + \boldsymbol{W}(k) \cdot \boldsymbol{u}_{in}(k) \tag{7}$$

Faster learning can be reached by higher learning rate α, however too high learning rate could lead to instability and errors. For stable learning process the learning rate should be lower than reciprocal value of highest eigen value of correlation matrix $\boldsymbol{u}(k)^T\boldsymbol{u}(k)$ of the input vector.

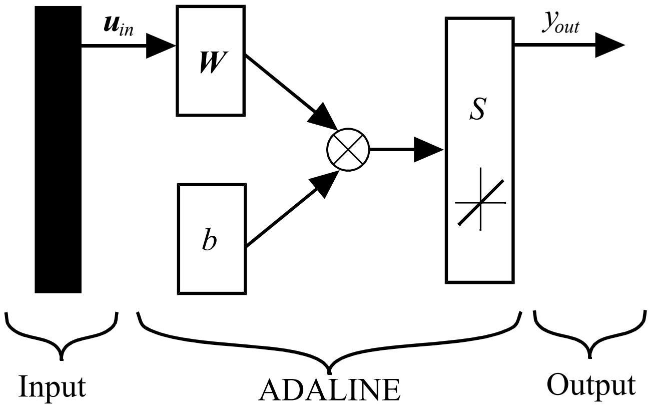

Fig.2 Simplified scheme of ADALINE

3. Multilayer feed-forward neural networks

Multilayer feed-forward neural networks were derived by generalization from Rosenblatt's perceptron, thus they are often called multilayer perceptrons (MLP). This type of artificial neural networks uses supervised training. One of the most known methods of supervised training is backpropagation algorithm; hence these ANNs are sometimes also called backpropagation networks.

In the MFFNN the signals flow between the neurones only in the forward direction i.e. towards the output. Neurones in MFFNN are organized in layers and neurones of the certain layer can have inputs from any neurones of the earlier layer. The ability to predict of ANN is determined by capability of modelling of certain process. By applying the Kolmogorov theorem it was proved that for general function approximation is sufficient two-layer MFFNN (one hidden layer) if non-polynomial transfer functions are used and the hidden layer has enough neurons [12].

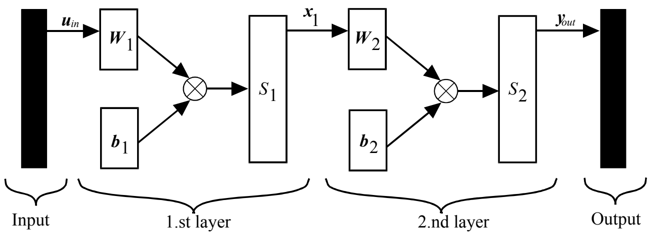

Fig.3 Simplified scheme of 2-layered MFFNN

The two-layered MFFNN, which contains one output layer and one hidden layers, is depicted in the figure 3 (this structure is implemented in this paper). This MFFNN can be described by the two equations:

$$\boldsymbol{y}_{out} = S_2\left(\boldsymbol{b}_2 + \boldsymbol{W}_2 \cdot \boldsymbol{x}_1\right) \tag{8}$$

$$\boldsymbol{x}_1 = S_1\left(\boldsymbol{b}_1 + \boldsymbol{W}_1 \cdot \boldsymbol{u}_{in}\right) \tag{9}$$

where $\boldsymbol{y}_{out}$ is the network output vector, S_i is transfer function of i-th layer, $\boldsymbol{b}_i$ is bias vector of i-th layer, $\boldsymbol{W}_i$ is weighting matrix of i-th layer, $\boldsymbol{x}_i$ is output vector of i-th layer and $\boldsymbol{u}_{in}$ is the network input vector.

4. Two interconnected cylindrical tanks

Let us consider SISO nonlinear system to be con-trolled which is shown in the figure 4 and consist two connected cylindrical tanks for liquid. The dynamic model if the system is (considering usual simplifications):

$$\frac{\pi d_1^2}{4}\frac{dh_1}{dt} + q_1 = q_{1v} \tag{10}$$

$$\frac{\pi d_2^2}{4}\frac{dh_2}{dt} + q_2 = q_1 \tag{11}$$

$$q_1 = k_1\sqrt{h_1 - h_2} \tag{12}$$

$$q_2 = k_2\sqrt{h_2} \tag{13}$$

where d_j is the diameter, h_j is liquid level and q_j is the output flow of the j-th tank. q_{1v} is the input flow to the first tank. The constants associated with the properties of pipes and valves are k_1 and k_2. The objective of the controller is to maintain the level in the first tank h_1 by adjusting the input flow q_{1v}.

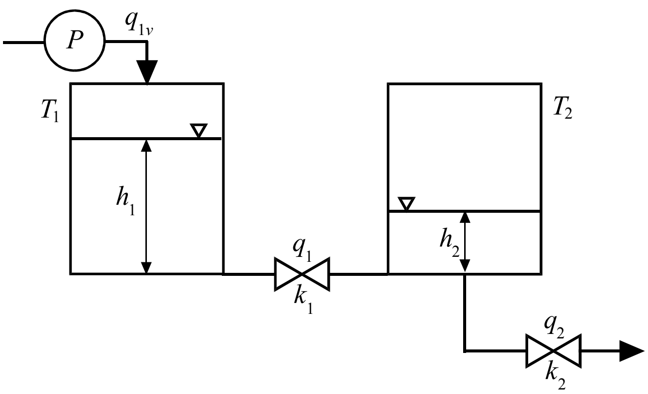

Fig.4 Scheme of two interconnected cylindrical tanks

This system is based on real-time laboratory model Amira DTS-200 (figure 5) which consist 3 cylindrical tanks and two pumps. However, in this contribution only two tanks (T_1 and T_2) and one pump were selected. Thus, the valves V_2 and V_4 were fully closed and the valve number 5 was set to the half position.

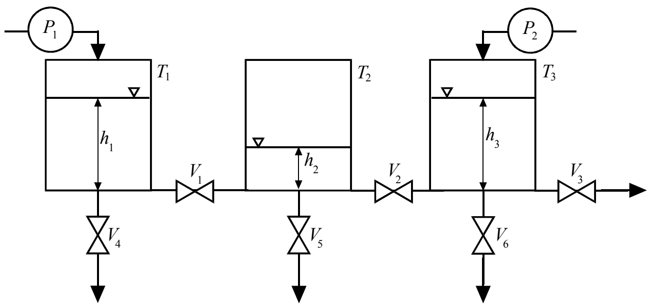

Fig.5 Scheme of DTS-200

The values of parameters are k_1 = 11.53 m$^{2.5}$/s and k_2 = 13.09 m$^{2.5}$/s were obtained by experiments on the DTS-200. The maximum input flow q_{1v} is 100 cm^3/s. The height of the tanks is 60 cm and their diameter is 14 cm.

The mathematical model of the system was programmed in the Simulink as S-Function. The Simulink scheme of the control loop is illustrated in figure 6.

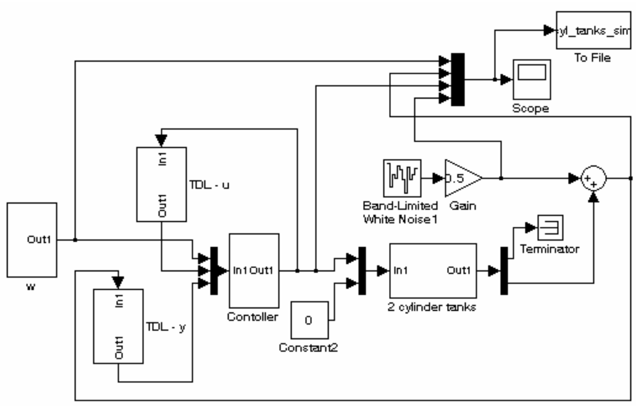

Fig.6 The used Simulink model of the control loop

5. Predictors based on artificial neural networks

In this paper two different artificial neural networks are applied, Adaptive linear network (ADALINE) and Multilayer feed-forward neural network (MFFNN). The first is used with on-line adaptation while the second one is using only off-line identification.

Nevertheless, in spite of ADALINE's ability to adapt we used off-line identification too in order to get rid of starting inaccuracies. However, the predictor was adapted at each sampling period so as to improve predictions. For the off-line identification was used input-output data generated by pulses of random amplitude and duration. Duration and amplitude of the pulses must be chosen carefully to produce accurate identification. We have used amplitudes in range <0; 100> cm^3/s and duration from 1s to 50s.

The MFFNN with 1 hidden layer was chosen the as a comparative method for prediction. The structure was 10→8→1. In the hidden layer the hyperbolic tangent function was used while in the output layer the linear transfer function was utilised. Because of the long training times of multilayer feed-forward neural networks, the off-line identification was necessary. Training data were same as for ADALINE and Levenberg-Marquart method was utilised as the training algorithm.

6. Simulations and results

Simulations were done for two types of predictor – adaptive linear network and multilayer feed-forward neural network. As can be seen from figure 6, the Band-Limited White Noise block was included to find out the behaviour without noise and while noise is pre-sent.

The gain of noise was set to 0 and 0.5 respectively. The sampling period of Simulink was set to 1s.

Both predictors were tested in control of the Simulink model of the two interconnected cylindrical tanks system in case of no noise and with presence of noise. In the following text are the simulations denoted as ADALINE1 (without noise), ADALINE2 (with noise), MFFNN1 (without noise) and MFFNN2 (with noise).

Due to constraints and nonlinear nature of predictors numerical optimization of the MPC criterion was necessary. The controller used constrained quasi-Newton method from Matlab Optimization Toolbox as a nonlinear optimization algorithm.

TDL blocks in the figure 6 represent so called Tapped Delay Line which stores five recent values of the signals. This block is necessary because ANN uses five recent values of system input and output. Example of the TDL is depicted in figure 7.

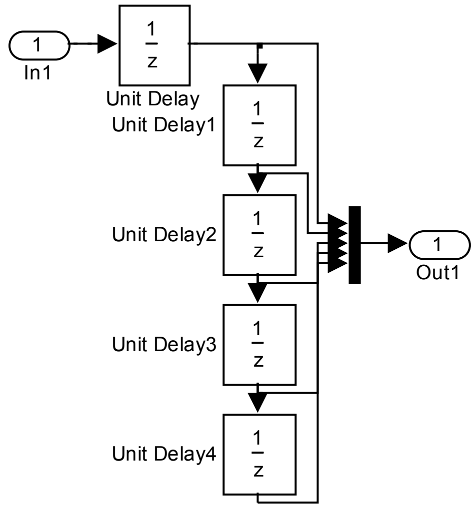

Fig.7 TDL block from the Simulink control loop

The parameters of the controller that used ADALINE based predictor were set in the following way: λ = 1, ρ = 0.2, N_1 =1, N_2 = 7, N_u = 5 in case of no noise (ADALINE1) and λ = 1, ρ = 0.8, N_1 =1, N_2 = 7, N_u = 5 in case of influence of noise (ADALINE2).

Simulation results are presented in figures 8 and 9.

The parameters of the controller that used MFFNN based predictor were set as follows: λ = 1, ρ = 5, N_1 = 1, N_2 = 10, N_u = 5 in case of no noise (MFFNN1) and λ = 1, ρ = 1, N_1 = 1, N_2 = 5, N_u = 5 in case of influence of noise (MFFNN2).

Simulation results are depicted in figures 10 and 11.

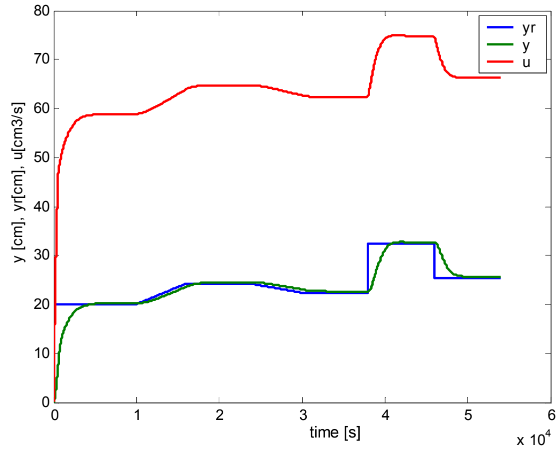

Fig.8 Simulation results of ADALINE1 (green – system output, blue – reference value, red - control signal)

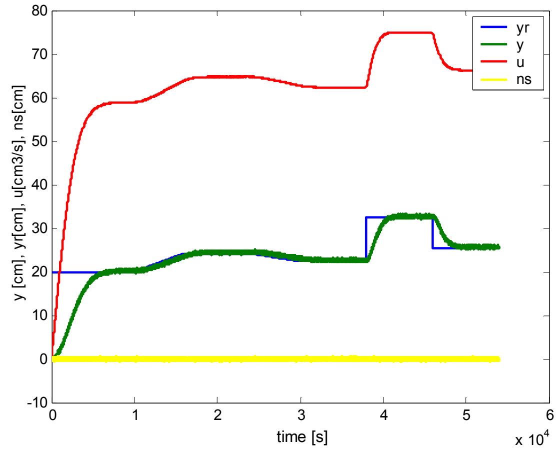

Fig.9 Simulation results of ADALINE2 (green – system output, blue – reference value, red - control signal, yellow - noise)

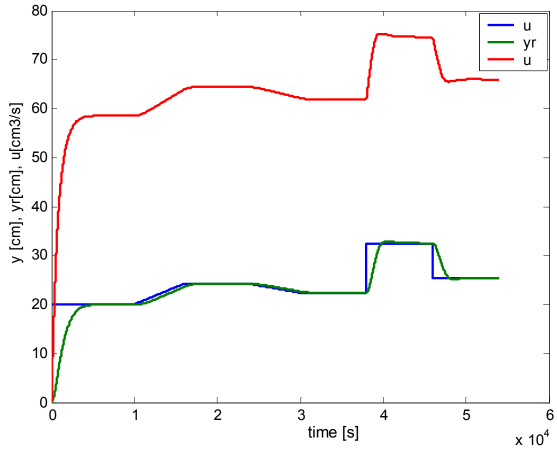

Fig.10 Simulation results of MFFNN1 (green – system output, blue – reference value, red - control signal)

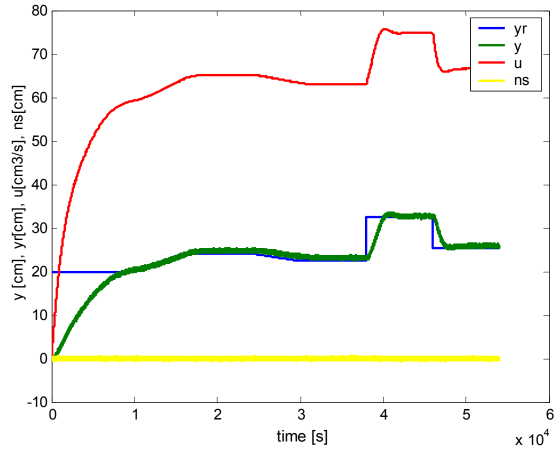

Fig.11 Simulation results of MFFNN2 (green – system output, blue – reference value, red - control signal, yellow - noise)

In order to compare results of both controllers (predictors) we used two quadratic criterions:

$$S_1 = \sum_k \left[y_r(k) - y(k) \right]^2 \qquad (14)$$

$$S_2 = \sum_k \left[y_r(k) - y(k) \right]^2 + \varphi \cdot \sum_k \left[u(k) - u(k-1) \right]^2 \qquad (15)$$

First criterion is based on control errors and represents the tracking performance of the controller. While the second criterion contains not only the control errors, but also the speed of control signal changes. Thus, it also represents the controller demands on the actuators. The φ parameter was se to 10000. The computed values of both criterions are shown in the table 1.

Simulation	S_1	S_2
ADALINE1	355177	400770
ADALINE2	912170	949451
MFFNN1	456429	480985
MFFNN2	1045880	1061729

Tab.1 Comparison of the simulations results

7. Discussion

As can be seen from figures 8 to 11, the usage of ADALINE as a predictor in model predictive control is possible for both cases - control without noise and with noise, despite the predictor was trained for data without noise. Due to influence of noise, the parameter ρ had to be increased to reduce the jittering of control actions. However, oscillations of output value could not be removed because of placement of noise at the output of the controlled system. Comparative method – MFFNN based predictor, provides similar results, however it has significantly longer training time (Off-line training of ADALINE lasted 0.1s but MFFNN took more than 10 minutes). Therefore, the simulation proved that simple one-neuron network with linear transfer function is able to predict the nonlinear system output with moderate deviations. Moreover, it was shown that the ADALINE can be used for sufficient predictive control of this kind of systems.

Conclusion

The main advantages of ADALINE are small memory requirements, fast training and simple usage. As a result of short training time ADALINE can be easily adapted on-line which increases the accuracy of control. Of course, the presented method has also disadvantages. The first drawback comes from the simplicity of ADALINE. Linear nature of ADALINE may result in not so exact predictions in comparison to predictions to be obtained from more complex ANN. On the other hand, MPC is quite tolerant to small predictor inaccuracy and on-line adaptation may also decrease the prediction error. The second disadvantage is the computational demands of optimization algorithm.

Acknowledgments

This work was supported by the Grant Agency of the Czech Republic under the grant 102/07/P137 and by the Ministry of Education, Youth and Sports of the Czech Republic under the grant MSM 7088352102. This support is gratefully acknowledged.

References

[1] CAMACHO, E. F., BORDONS, C.: Model Predictive Control in the Process Industry. Springer – Verlag, Berlin, 1995

[2] HAGAN, M., DEMUTH, H., JESUS, O. D.: An Introduction to the Use of Neural Networks in Control Systems. Int. Journal of Robust and Nonlinear Control. Vol. 12, No. 11, pp. 959-985, 2002

[3] KANJILAL, P.P.: Adaptive prediction and predictive control. Peter Perengrinus Ltd., London, 1995

[4] LIU, G.P.: Nonlinear identification and control – A neural network Approach. Springer - Verlag London Ltd., London, 2001

[5] MESZAROS, A., RUSNAK, A., FIKAR, M.: Adaptive Neural PID Control – Case Study: Tubular Chemical Reactor. Computers and Chemical Engineering Supplement. Vol. 1999, pp. 847-850, 1999

[6] LENNOX, B., MONTANGUE, G. A., FRITH, A. M., GENT, CH., BEVAN, V.: Industrial application of neural networks – an investigation. Journal of Process Control. Vol. 11, No. 5, pp. 497-507, 2001

[7] KAMBHAMPATI, C., MASON, J.D., WARWICK, K.: A stable one-step-ahead predictive control of nonlinear systems. Automatica. Vol. 36, No. 4, pp. 485-495, 2000

[8] YU, D.W., YU, D.L.: Neural network control of multivariable processes with fast optimization algorithm. Neural Computing & Applications, Vol. 12, No. 3-4, pp. 185-189, 2003

[9] WIDROW, B., HOFF, M. E. JR.: Adaptive switching circuits. IRE WESCON Convention Re-cord, Vol. 1960, No. 4, pp. 96–104, 1960

[10] ARBIB, M. A.: The Handbook of Brain Theory and Neural Networks. Second Edition. The MIT Press, Cambridge, Massachusetts, 2002.

[11] FREEMAN, J. A., SKAPURA D. M.: Neural Networks Algorithms, Applications, and Programming Techniques, Addison Wesley Publishing Company, Reading, MA, 1991

[12] LESHNO, M., LIN, V. Y. PINKUS, A. SCHOCKEN, S.: Multilayer feedforward networks with a nonpolynomial activation can approximate any function. Neural networks. Vol. 6, No. 6, pp. 861-867, 1993

Ing. David Sámek, Ph.D.

Tomas Bata University in Zlín
Faculty of Technology
Department of Production Engineering
Mostní 5139
76001 Zlín
Tel.: +420 576 035 157
Fax: +420 576 035 176
E-mail: samek@ft.utb.cz

Prof. Ing. Petr Dostál, CSc.

Tomas Bata University in Zlín
Faculty of Applied Informatics
Department of Process Control
Mostní 5139
76001 Zlín
Tel.: +420 576 035 195
Fax: +420 576 035 279
E-mail: dostalp@fai.utb.cz