David SAMEK*, Petr DOSTAL**

ARTIFICIAL NEURAL NETWORK WITH RADIAL BASIS FUNCTION IN MODEL PREDICTIVE CONTROL OF CHEMICAL REACTOR

SUMMARY

This paper describes the application of artificial neural network with radial basis function as a predictor in model predictive control. Radial basis function neural networks are known for their fast training. Thus, this type of artificial neural networks offers promising way how to reduce computational cost during offline predictor training and eventual online adaptation. The features of this type of artificial neural network are presented in simulations in MATLAB/Simulink on the nonlinear system control. The aim of this paper is to suggest one approach how to solve nonlinear prediction problem using artificial neural network respecting computational demands of the predictor.

***Keywords**: artificial neural network, radial basis function, model predictive control*

SZTUCZNE SIECI NEURONOWE Z RADIALNYMI FUNKCJAMI BAZOWYMI W PREDYKTYWNYM STEROWANIU REAKTOREM CHEMICZNYM

Artykuł jest poświęcony zastosowaniu sztucznych sieci neuronowych z radialnymi funkcjami bazowymi jako predykatora w modelach sterowania predyktywnego. Sieci radialne są znane z możliwości ich szybkiego uczenia. Dlatego ten typ sztucznych sieci neuronowych umożliwia redukcję zasu obliczeń podczas uczenia sieci w trybie off-line i ewentualnych zastosowań on-line. Cechy omawianych aplikacji sieci neuronowych przedstawiono w symulacyjnych obliczeniach sterowania nieliniowego układu z wykorzystaniem środowiska MATLAB/Simulink.

***Słowa kluczowe**: sztuczne sieci neuronowe, radialne funkcje bazowe, sterowanie predykcyjne*

1. INTRODUCTION

The predictor in model predictive controller (MPC) plays a key role. Selection of suitable model is the first and most important question in this control method. The classic approach to prediction is based on a set-upping of mathematical model of system. The disadvantage of this approach is the necessity to know system, its parameters and to describe them. Moreover, difficulties rise during the prediction of strongly nonlinear systems and the eventual control performance is not satisfactory. Artificial neural networks offer interesting possibility for modelling and predicting any nonlinear process output without a priori knowledge. Artificial neural networks (ANN) can be regarded as nonlinear black-box models. The most used type of artificial neural networks for predictive control is multilayer-feed forward neural network (MFFNN). This type of artificial neural network suffers many problems, such as the long training times, complex structure, memory demands and a lot of parameters, which have key influence to the particular results. Nevertheless, there are many other artificial neural networks that are suitable for nonlinear system prediction, for example artificial neural networks with radial basis function.

In this paper the application of radial basis function artificial neural network in MPC controller is presented, while as a comparison method the multilayer feed-forward neural network is used.

2. RADIAL BASIS FUNCTION NEURAL NETWORKS

Radial Basis Function (RBF) networks are actually feed-forward networks with one layer consisting local units. These networks are significantly younger than perceptron and perceptron-based networks. Nevertheless, theoretical and practical results showed that RBF networks represent interesting alternative for classic MFFNNs. Training of RBF network is much faster than training of MFFNN. On the other hand, RBF networks usually require more neurons for same problem and they cannot treat non-relevant inputs. Because of their worse generalization, RBF networks perform the best if large amount of training data is used. RBF networks are mainly applied for approximation and function interpolation.

The RBF network is a single-hidden-layer feed-forward network with linear output transfer functions S_2 and nonlinear transfer functions S_1 on the hidden layer nodes. Many types of nonlinearities may be used. There is also typically a bias on each output node. The primary adjustable parameters (see Fig. 1) are the output layer weights w_2, weights w_1 connecting the nodes from input layer with the nodes from the hidden layer. The hidden layer weights w_1 can be regarded as representing "prototypes" of patterns in the input space, either as cluster centres or as pattern exemplars that are in some sense representative of the distribution of input patterns. Similarly, the hidden layer weights b_1, that govern the regularization of the network (after the model order complexity of the number of basis functions), are sometimes considered to represent the range of influence of these prototypes (though this intuition can break down for some of the nonlocal basis functions that can be employed) (Lowe 2002).

* Department of Production Engineering, Faculty of Technology, Tomas Bata University in Zlin, Zlin, Czech Republic, samek@ft.utb.cz
** Department of Process Control, Faculty of Applied Informatics, Tomas Bata University in Zlin, Zlin, Czech Republic, dostalp@fai.utb.cz

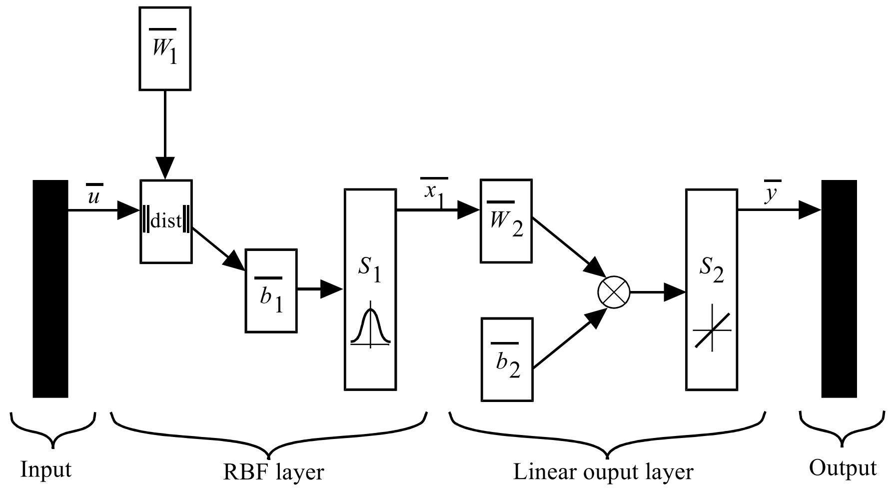

Fig. 1. Structure of RBF network

The mathematical embodiment of the RBF neural network takes the following form (using simplified matrix notation).

$$\mathbf{y_{out}} = S_2\left(\mathbf{b_2} + \mathbf{W_2} \cdot \mathbf{x_1}\right) \tag{1}$$

$$\mathbf{x_1} = S_1\left(\left\|\mathbf{W_1} - \mathbf{u_{in}}\right\| \cdot \mathbf{b_1}\right) \tag{2}$$

where $\mathbf{y_{out}}$ is the vector of the output data, $\mathbf{u_{in}}$ is the input vector, $\mathbf{W_i}$ and $\mathbf{b_i}$ denotes weighting matrix and bias vector of i-th layer, S_i is transfer function of i-th layer and $\mathbf{x_1}$ stands for the hidden layer output vector.

3. SIMULATIONS AND RESULTS

In this chapter the system to be controlled is introduced. Then, the predictors and the controller are designed. Final part of this chapter presents simulations results.

3.1. Continuous stirred tank reactor

In the chemistry industry the continuous stirred tank reactors (CSTR) are commonly used. The control of this kind of reactor is one of typical process control tasks and a lot of papers have been published about this topic for example (Bravo *et al.* 2006; Nikravesh *et al.* 2000; Sitsu and Bequette 1991).

Let us consider the nonlinear single input – single output system of continuous stirred tank reactor (CSTR) which can be described as follows (Brengel and Seider 1989; Hagan *et al.* 2002; Li and Biegler 1988):

$$r_b = \frac{K_1 \cdot C_b}{\left(1 + K_2 \cdot C_b\right)^2} \tag{3}$$

$$\frac{dh}{dt} = q_1 + q_2 - 0.2\sqrt{h} \tag{4}$$

$$\frac{dC_b}{dt} = \left(C_{b1} - C_b\right)\frac{q_1}{h} + \left(C_{b2} - C_b\right)\frac{q_2}{h} - r_b \tag{5}$$

where r_b is rate of consumption of C_b, h is the liquid level, C_b is the product concentration at the output of the process, q_1 is the flow rate of the concentrated feed C_{b1}, and q_2 is the flow rate of the concentrated feed C_{b2}. The input concentrations are set to C_{b1} = 24.9 mol/cm^3 and C_{b2} = 0.1 mol/cm^3 (see Fig. 2). The constants associated with the rate of consumption are K_1 = K_2 = 1. The task of the controller is to control the product concentration C_b by adjusting the flow rate q_1. According to (Brengel and Seider 1989; Hagan *et al.* 2002) it assumed that parameter q_2 is constant (q_2 = 0.1 cm^3/s) and the level of liquid in the reactor is not controlled. Thus, the system is regarded as single input – single output (SISO). The allowable range for q_1 was assigned to be in the interval <0, 4>. The initial conditions are $C_b(0)$ = 22 mol/cm^3 and $h(0)$ = 30 cm.

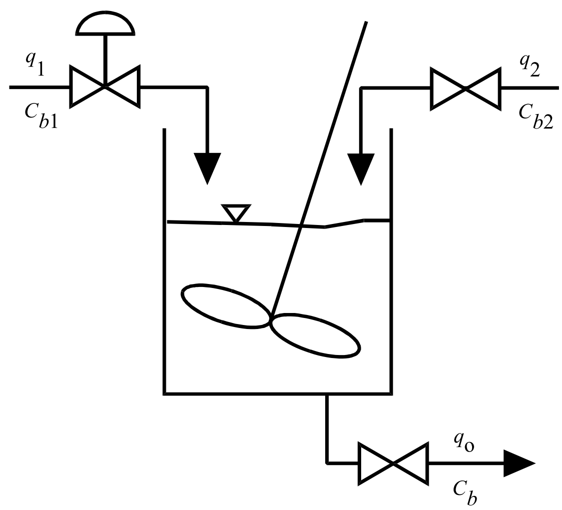

Fig. 2. The continuous stirred tank reactor

3.2. Predictor

All tested predictors were trained off-line using identification data generated by stepped actuating signal. For the off-line identification were 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. The amplitudes in range <0; 4> cm^3/s and duration from 1 s to 50 s were used. For the predictions five last input and output values were used (predictors had 10 inputs).

The predictor based on RBF network consisted of 344 neurons in hidden layer and 1 neuron in the output layer. The predictor based on multilayered feed-forward neural network consisted of 50 neurons with hyperbolic tangent transfer function in the hidden layer and 1 neuron with linear transfer function in the output layer. The average training time for the predictor based on the multilayer feed-forward neural network was 20.3 min. However, the average training time for the predictor based on the RBF networks was only 3.0 min. All the predictors were designed using Matlab Neural Network Toolbox.

3.3. Controller

There are various approaches to predictive control by artificial neural networks. Generally, we can say that these methods use ANN as the process model in order to get its output predictions. The most often used approach is model predictive control (Camacho and Bordons 2007). MPC is a broad control strategy applicable to both linear and non-linear 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:

$$J = \lambda \cdot \sum_{i=N_1}^{N_2} \left[w(k+i) - \hat{y}(k+i) \right]^2 + \rho \cdot \sum_{i=1}^{N_u} \left[u_t(k+i-1) - u_t(k+i-2) \right]^2 \tag{6}$$

where N_1, N_2 and N_u define horizons over which the tracking error and the control increments are evaluated (usually $N_2 = N_u$). The u_t variable is the tentative control signal, w is the desired response and $\hat{y}$ is the network model response. The parameters λ and ρ determine the contribution that the sums of the squares of the future control errors and control increments have 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. The MPC strategy using receding horizon is depicted in Figure 3.

There is usually assumed that after a certain interval $N_u < N_2$ there is no variation in the proposed control signals, that is:

$$\Delta u(k+i) = 0 \quad \text{for} \quad i \in \langle N_u, N_2 - 1 \rangle \tag{7}$$

This is equivalent to giving infinite weights to the changes in the control from a certain instant. This approach is applied in this paper.

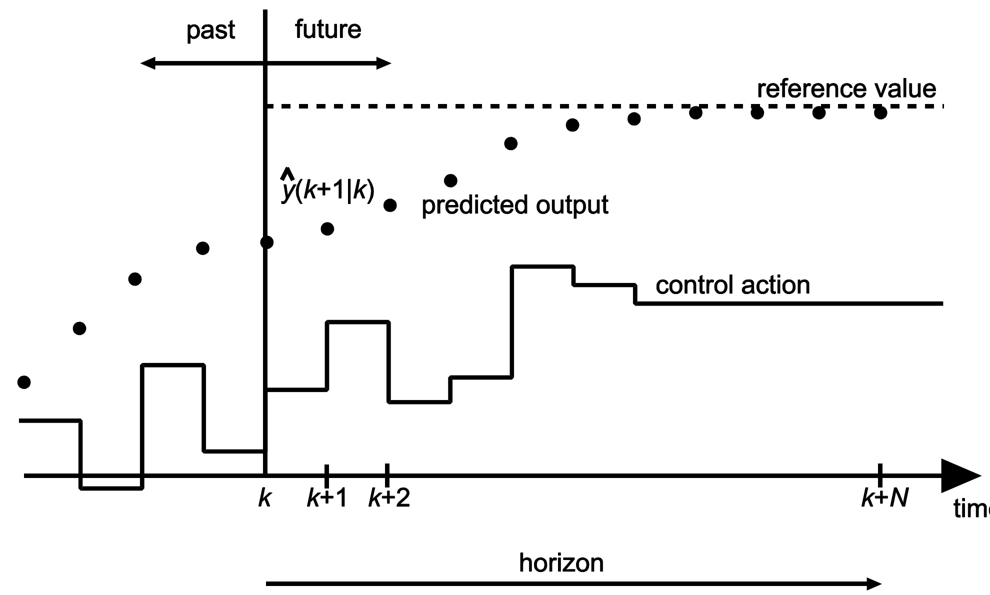

Fig. 3. MPC strategy – receding horizon

Due to constraints and nonlinear nature of predictors the numerical optimization of the MPC criterion (6) was necessary. The controller used constrained quasi-Newton method from Matlab Optimization Toolbox as a nonlinear optimization algorithm. The sampling time was 1s according to (Brengel and Seider 1989; Li and Biegler 1988).

3.4. Results

The predictors were tested in the hereinabove described MPC controller, while various settings of the controller were used. We have tested a lot of various controller settings, but in this paper only one simulation for each predictor is presented. Results are depicted in Figures 4, 5 and the Table 1.

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

$$S_y = \sum_k (w_k - y_k)^2 \tag{8}$$

$$S_u = \sum_k (u_k - u_{k-1})^2 \tag{9}$$

The first criterion S_y is based on control errors and represents the tracking performance of the controller. While the second criterion S_u is based on the changes of the control signal and it represents the controller demands on the actuators.

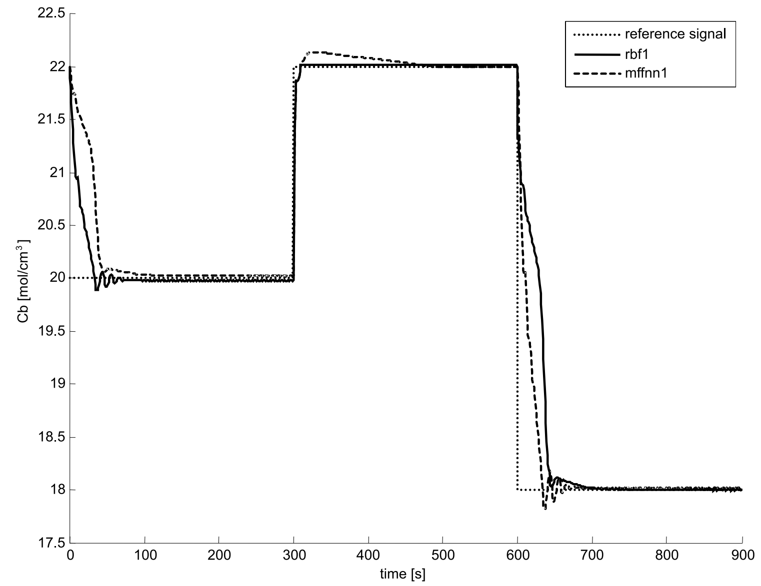

Fig. 4. System response

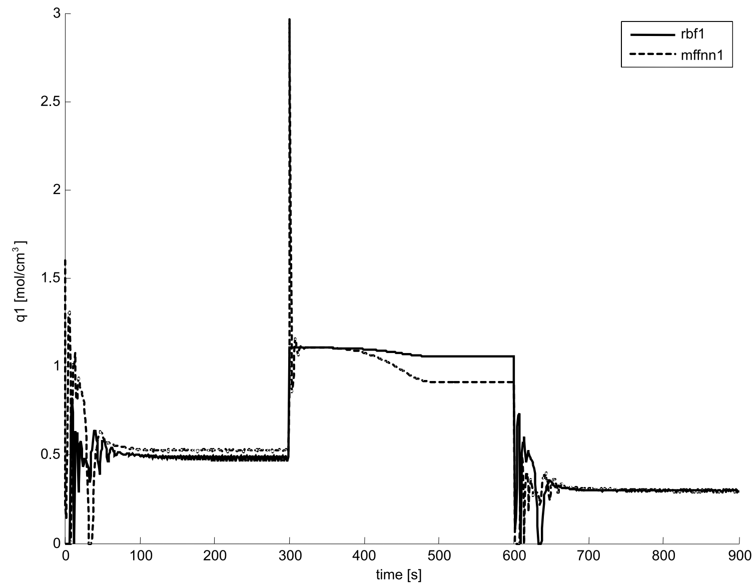

Fig. 5. Control signal

Table 1

Comparison of the controllers

Predictor	N_1	N_2	N_u	λ	ρ	S_y	S_u
RBF	1	5	5	1	0.1	262.6	2.8
MFFNN	1	5	3	1	0.01	211.6	11.8

4. CONCLUSIONS

As can be seen from the simulations results, both types of artificial neural networks – MFFNN and RBF networks proved their qualities. The control signal produced by the MPC controller with MFFNN predictor is much more jittering and oscillating then the one produced using the RBF network. This behaviour causes higher overshoots of the controlled system output after step change of reference signal. From the point of view of the S_u criterion the radial basis function neural network based controller provides better results. On the other hand, the criterion S_y shows lower value for the multilayer feed-forward neural network. Thus, it could be stated that control quality criterions are more or less comparable for both artificial neural networks. However, the training times of the artificial neural network that uses radial basis function are dramatically shorter.

The disadvantage of the radial basis function neural networks usage is the higher memory requirements, which is caused by the different neural network training system in comparison to multilayer feed-forward neural networks. The number of neurons in the hidden layer of RBF network is increased at each training epoch till the training goal (training error) is not reached. Nevertheless, memory capacity of modern computers and consequently controllers nowadays offers enough space for middle-sized artificial neural networks.

It can be concluded that radial basis function neural networks may be suggested as another possible way how to cope with computational demands of ANN based predictors in model predictive controllers, especially in case of need of an on-line adaptation of the predictor during the control. Application of this methodology in adaptive predictive controllers will be subject of further research.

Acknowledgement

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

References

Bravo J.M., Alamo T., Camacho E.F. 2006: *Robust MPC of constrained discrete-time nonlinear systems based on approximated reachable sets*. Automatica, vol. 42, no. 10, pp. 1745–1751.

Brengel D.D., Seider W.D. 1989: *Multistep Nonlinear Predictive Controller*. Industrial and Engineering Chemistry Research, vol. 28, no. 12, pp. 1812–1822.

Camacho E.F., Bordons C. 2007: *Model Predictive Control – Second edition*. Springer – Verlag, London, 405p, ISBN 978-1-85233-694-3.

Hagan M., Demuth H., Jesus O.D. 2002: *An Introduction to the Use of Neural Networks in Control Systems*. International Journal of Robust and Nonlinear Control, vol. 12, no. 11, pp. 959–985.

Li W.C., Biegler L.T. 1988: *Process Control Strategies for Constrained Non-linear Systems*. Ind. Eng. Chem. Res., vol. 27, no. 8, pp. 1421–1433.

Lowe D. 2002: *Radial basis function networks*. in: M.A. Arbib (Ed.), The Handbook of Brain theory and Neural Networks, MIT Press, Cambridge, MA, ISBN 0-262-01197-2, pp. 779–782.

Nikravesh M., Farell A.E., Stanford T.G. 2000: *Control of nonisothermal CSTR with time varying parameters via dynamic neural network control (DNNC)*. Chemical Engineering Journal, vol. 76, no. 1, pp. 1–16.

Sitsu P.B., Bequette B.W. 1991: *Non-linear predictive control of uncertain processes: application to CSTR*. American Institute of Chemical Engineers Journal, vol. 37, no. 11, pp. 1711–1723.