Global models of dynamic complex systems – modelling using the multilayer neural networks

Grzegorz Drałus[*]

*Faculty of Electrical and Computer Engineering, Rzeszów University of Technology,
W. Pola 2, 35-959 Rzeszów, Poland*

Abstract

In this paper, global models of dynamic complex systems using the neural networks is discussed. The description of a complex system is given by a description of each system element and structure. As a model the multilayer neural networks with the tapped delay line (TDL), which have the same structure as a complex system, are accepted. Two approaches, a global model and a global model with the quality local model taken into account are proposed.

To learn global models the modified back-propagation algorithms have been developed for the unique structure of the complex model. To model dynamic simple plants, of which the complex system is composed, a series-parallel model of identification using the feedforward network with the tapped delay line (TDL) and the feedback loops, in which the gradient can be calculated by means of the simpler static back-propagation method is proposed. Computer simulations were performed for the dynamic complex system, which consists of two dynamic nonlinear simple plants connected in series, described by means of nonlinear difference equations.

1. Introduction

There are various of kinds of complex systems. The first kind is composed of blocks (simple plants), consisting of identification of particular process blocks and their interconnections [1]. The other one is a collection of systems that are not approximable (arbitrarily close) by a finitely parameterized collection of systems [2]. Another example are the complex systems as the complex operation ones [3].

In this paper the dynamic complex system, is a system which consists of dynamic simple plants connected in the internal structure of the system, that means the flow of signals between particular simple plants. Simple plant means the input-output system.

---

[*]*E-mail address*: gregor@prz.rzeszow.pl

Modelling methods of dynamic systems are well known and extensively represented in literature. However, modelling of dynamic complex systems is a very difficult problem and not solved so far. For example, in the complex control of industry processes it is necessary to specify the control algorithm for different phases of the process. Developing models of complex systems is forced by the need for complex control of complex industrial systems [1].

Mathematical methods of the complex system identification are usually based on the modelling of simple plants being parts of complex systems and on their optimisation. Then, the model of a complex system consisting of optimal simple models is built up. However, such models are not global models [1].

The main goal of this paper is to show that it is possible to achieve global models of dynamic complex systems using the multilayer neural networks with the feedback loop.

One of typical complex systems is the collection series connecting simple plants without the feedback loop. In this structure the output preview of a simple plant is the input to the next simple plant. So, the global model of the complex system is obtained by connecting simple models in the same series (cascade) structure. Output signals of simple models and simple plants, and output signals of the global model and the complex system will be used to construct a quality criterion of the global model. Two cases will be taken into account, depending on the way of considering a particular simple model in the global model. The first case is the global model, in which signals flow during the learning mode and working modes are the same as in the complex system. The global model has the same structure as the complex system. For this model the global quality criterion was defined. Adjusting parameters for this model (learning of the net) should be done in such a way that the model outputs are close to the complex system outputs, in the same structure. The second task is the difference between the simple model outputs and the simple plant outputs should be small.

The second case is the global model with the quality of local models taken into account. The quality of local models in a global model can be considered by the acceptance of the synthetic criterion. This criterion assures that the global model will be satisfactory and local models which the global model consists of, will have sufficiently good quality. Both models have the same structure but they differ in the way of learning.

2. Models of complex systems

For the modeling of simple dynamic plants with multilayer neural networks, usually the series-parallel model is used [6]. In a such structure, present and past input and output signals transmitted via TDL are used. In this case there is no feedback from the network, but from the plant. Then, the neural network works

as an unidirectional one and it can be learned according to the "delta" rule (the static backpropagation algorithm). After the learning phase (stage), if the error of modeling is sufficiently small, the series-parallel model can be replaced by a parallel model with the feedback loop from the model outputs without significant consequences [6].

2.1. The global model of the dynamic complex system (GM)

The dynamic complex system can consist of R simple plants. An example of a complex system which consists of two simple plants ($R = 2$) and its global model is shown in Fig. 1. The global model which consists of two simple models (noted as M_1 i M_2), is connected in the structure corresponding to this system. In such a structure the output preview simple model is the input to the next simple model. The outputs of a global model are described:

$$\hat{\mathbf{y}}^{(r)} = f_r\left(\hat{\mathbf{y}}^{(r-1)}, \mathbf{w}^{(r)}\right) = f_r\left(f_{r-1}\left(\cdots f_1\left(\mathbf{u}^{(1)}, \mathbf{w}^{(1)}\right), \cdots, \mathbf{w}^{(r-1)}\right), \mathbf{w}^{(r)}\right),$$
$$= \hat{f}_r\left(\mathbf{u}^{(1)}, \mathbf{w}^{(1)}, \ldots, \mathbf{w}^{(r)}\right) \tag{1}$$

where: $\mathbf{u}^{(1)}$ is the external input for the global model and the complex system, $\mathbf{w}^{(r)}$ are the parameters (weights) of r-th simple model.

Passing on the input of the model k-th sample of input vector $\mathbf{u}^{(1)}(k)$ we get the output signals of the r-th simple model $\hat{\mathbf{y}}^{(r)}(k) = [\hat{y}_1^{(r)}(k), \hat{y}_2^{(r)}(k), \cdots, \hat{y}_{J_r}^{(r)}(k)]$, where J_r is the number of outputs of the r-th simple model.

The error for k-th learning data sample for the r-th simple model is as follows:

$$\hat{\mathbf{e}}^{(r)}(k) = \hat{\mathbf{y}}^{(r)}(k+r) - \mathbf{y}^{(r)}(k+r). \tag{2}$$

The global quality assessment criterion for this global model is the sum of squared errors (differences) between the outputs of simple models (M_r) and the outputs of respective simple plants (P_r) of the complex system:

$$Q_g(\mathbf{W}) = \sum_{r=1}^{R} \beta_r Q_g^{(r)}\left(\mathbf{w}^{(r)}\right) = \frac{1}{2}\sum_{k=1}^{K}\left(\hat{\mathbf{e}}(k)\right)^{\mathrm{T}} \boldsymbol{\beta}\, \hat{\mathbf{e}}(k), \tag{3}$$

and the quality assessment criterion of a simple model in the global model:

$$Q_g^{(r)}\left(\mathbf{w}^{(r)}\right) = \frac{1}{2}\sum_{k=1}^{K}\sum_{j=1}^{J_r}\left(\hat{y}_j^{(r)}(k+r) - y_j^{(r)}(k+r)\right)^2, \tag{4}$$

where: $\boldsymbol{\beta} = diag[\beta_1, \beta_2, \ldots, \beta_R]$, $\beta_r \in [0,1]$, $r = 1, \ldots, R$ and $\sum_{r=1}^{R}\beta_r = 1$, K – the length of input vector, $\mathbf{W} = \{\mathbf{w}^{(1)}, \mathbf{w}^{(2)}, \ldots, \mathbf{w}^{(R)}\}$ – is a set of parameters (weights) of the global model divided into subsets of simple models.
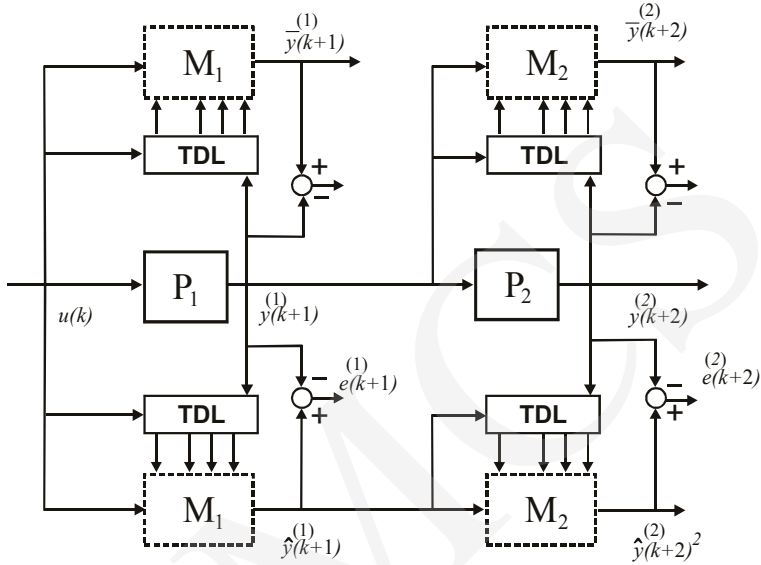
Fig. 1. The two-parts complex system and their global models

The coefficients β_r determine the effect of individual simple models on the global quality criterion.

As a global model the multilayer neural networks with TDL were accepted. To determine these optimal model parameters **W** the modified backpropagation learning algorithm for the dynamic complex systems presented in detail in [5] can be used.

2.2. The global model for dynamic complex system with the quality local models taken into account

In the global model with the quality local models taken into account the compromise between the global model quality and local models quality is to be found. Parameters of such a model should ensure that the global model is satisfactory and local models are of sufficiently good quality.

For the complex system shown in Fig. 1 there is built the global model which has the same structure. In the global model there can be pointed out the local models, which are denoted as M_1 and M_2. The outputs $\overline{\mathbf{y}}^{(r)}$ of the local models are described by the formula:

$$\overline{\mathbf{y}}^{(r)} = \overline{f}_r\left(\mathbf{u}^{(r)}, \mathbf{w}^{(r)}\right) \ , \tag{5}$$

where the input of the *r*-th local model $\boldsymbol{u}^{(r)} = \boldsymbol{y}^{(r-1)}$ is the output of the previous simple plant.

The error of the *r*-th local model for the *k*-th learning data sample is as follows:

$$\overline{\mathbf{e}}^{(r)}(k+r) = \overline{\mathbf{y}}^{(r)}(k+r) - \mathbf{y}^{(r)}(k+r). \tag{6}$$

The differences between the outputs of local models (M_r) and those of respective simple plants (P_r) of the complex system are used to define the quality criterion of local models :

$$Q_r\left(\mathbf{w}^{(r)}\right) = \frac{1}{2}\sum_{k=1}^{K}\left(\overline{\mathbf{e}}^{(r)}(k)\right)^{\mathrm{T}}\overline{\mathbf{e}}^{(r)}(k) = \frac{1}{2}\sum_{k=1}^{K}\sum_{j=1}^{J_r}\left(\overline{y}_j^{(r)}(k+r) - y_j^{(r)}(k+r)\right)^2. \tag{7}$$

For the global model with respect to the quality of local models, the synthetic criterion was accepted:

$$Q_w(\mathbf{W}) = \alpha_0 \; Q_g(\mathbf{W}) + \alpha_1 \; Q_1\left(\mathbf{w}^{(1)}\right) + \cdots + \alpha_r \; Q_r\left(\mathbf{w}^{(r)}\right) + \cdots + \alpha_R \; Q_R\left(\mathbf{w}^{(R)}\right), \tag{8}$$

where: $0 \le \alpha_r \le 1$ for $r = 0,1,\ldots,R$, $\sum_{r=0}^{R}\alpha_r = 1$.

The coefficients $\boldsymbol{\alpha} = [\alpha_1 \; \alpha_2 \; \ldots \; \alpha_R]$ determine the effect of individual local models on the synthetic quality criterion (8). The coefficient α_0 determines the effect of global quality criterion (3) on the synthetic quality criterion (8). By suitable choice of α_r coefficients the influence of the local models on the global model can be determined.

2.3. Modified backpropagation learning algorithm for the global model with quality local models taken into account (GLM)

For learning the multilayer neural networks, of which the GLM model is constructed, it was necessary to modify the well-known gradient error backpropagation algorithm. This modification must consider the situation that a complex model consists of the static network with the tapped delay line and the feedback loop form a complex system during the learning mode and the feedback loop forms the complex model when the model works.

The synthetic quality criterion (8) for the *k*-th data learning sample can be written as:

$$Q_w^k(\mathbf{W}) = \alpha_0 \beta_R \frac{1}{2}\sum_{j=1}^{J_r}\left(\hat{y}_j^{(R)}(k+R) - y_j^{(R),k}(k+R)\right)^2$$
$$+ \frac{1}{2}\sum_{r=1}^{R}\alpha_r \sum_{j=1}^{J_r}\left(\overline{y}_j^{(r)}(k+r) - y_j^{(r)}(k+r)\right)^2. \tag{9}$$

In criterion (9) (model GLM) all coefficients β_r are equal zero, except for the last one $\beta_R = 1$. In order to minimise the error function $Q_w^k(\mathbf{W})$, the gradient descend method can be now applied. Using this method for the neural network learning, one obtains the following weight increment:

$$\Delta w_{ji}(k) = -\eta \frac{\partial Q_w^k(\mathbf{W})}{\partial w_{ji}} = -\eta \frac{\partial Q_w^k(\mathbf{W})}{\partial z_j^m} \frac{\partial z_j^m}{\partial w_{ji}}, \quad (10)$$

when: η – the learning rate from the range (0.1).

Let us define the „delta" error $\delta_j^{(r),m,k}$ arising in j-th element of m-th layer in r-th element of the model for k-th sample:

$$\delta_j^{(r),m,k} = \frac{\partial Q_w^k(\mathbf{W})}{\partial z_j^m(k)} = \frac{\partial Q_w^k(\mathbf{W})}{\partial u_j^m(k)} \frac{\partial u_j^m(k)}{\partial z_j^m(k)} = \frac{\partial Q_w^k(\mathbf{W})}{\partial u_j^m(k)} f'\left(z_j^m(k)\right), \quad (11)$$

where $\tilde{u}_j^m(k)$ is the output signal of j-th neuron of m-th layer of network, $\tilde{u}_j^m(k)$
$= \hat{u}_j^m(k)$ in the global model and $\tilde{u}_j^m(k) = \overline{u}_j^m(k)$ in local models.

At this point, one has to introduce the concept of the "additional" hidden layers. The "additional" layers are those whose outputs are also the outputs of sub-models.

Omitting the calculation details, the „delta" errors for particular layers can be calculated as follows:

– for the output layer,

$$\delta_j^{(R),M,k} = \alpha_0 f'\left(z_j^M(k)\right)\left(\hat{y}_j^{(R)}(k+R) - y_j^{(R)}(k+R)\right)$$
$$+ \alpha_R f'\left(z_j^M(k)\right)\left(\overline{y}_j^{(R)}(k+R) - y_j^{(R)}(k+R)\right) \quad (12)$$

– for the common hidden layer:

$$\delta_j^{(r),m,k} = f'\left(z_j^{m,k}\right)\sum_{l=1}^{I_{m+1}} \delta_l^{(r),m+1,k} \; w_{lj}^{(r),m+1}, \quad (13)$$

– for the "additional" hidden layer:

$$\delta_j^{(r),m,k} = f'\left(z_j^m(k)\right)\sum_{l=1}^{I_{m+1}} \delta_l^{(r+1),(m+1),k} \; w_{lj}^{(r+1),m+1}$$
$$+ \alpha_r f'\left(z_j^{m,k}\right)\left(\overline{y}_j^{(r)}(k+r) - y_j^{(r)}(k+r)\right) \quad (14)$$

With the above definitions, the general delta formula for weight increments takes the form:

$$\Delta w_{ji}^{(r),m} = -\eta \sum_{k=1}^K \delta_j^{(r),m,k} \; \hat{u}_i^{m-1}(k+r) - \eta \sum_{k=1}^K \delta_j^{(r),m,k} \; \overline{u}_i^{m-1}(k+r). \quad (15)$$

The *Rprop* algorithm [7], which does not use the information about the values of gradients allows for the significant increase in the speed of the learning process, especially in the areas where the error function gradient is not steep. The adjustment of the parameters in the l-th step of optimization in the modified *Rprop* method is:

$$\Delta w_{ji}^{(r),m} = \sum_{k=1}^{K} \Delta w_{ji}^{(r),m}(k) =$$

$$-\eta_{ji}^{l} \operatorname{sgn}\left(\sum_{k=1}^{K} \delta_{j}^{(r),m,k} \hat{u}_{i}^{m-1}(k+r)\right) -\eta_{ji}^{l} \operatorname{sgn}\left(\sum_{k=1}^{K} \delta_{j}^{(r),m,k} \bar{u}_{i}^{m-1}(k+r)\right) \tag{16}$$

where

$$\eta_{ji}^{l} = \begin{cases} \min\left(a \cdot \eta_{ji}^{(l-1)}, \eta_{max}^{l}\right) & \text{for} \quad S_{ji}^{l} \cdot S_{ji}^{(l-1)} > 0 \\ \max\left(b \cdot \eta_{ji}^{(l-1)}, \eta_{min}^{l}\right) & \text{for} \quad S_{ji}^{l} \cdot S_{ji}^{(l-1)} < 0 \\ \eta_{ji}^{(l-1)} & \text{for} \quad S_{ji}^{l} \cdot S_{ji}^{(l-1)} = 0 \end{cases} \tag{17}$$

(17)

and l – the number epoch, $S_{ji}^{l} = \partial Q_{w}(\mathbf{W}(l))/\partial w_{ji}$; a=1.2; b=0.5; $\eta_{max} = 50$; $\eta_{min} = 10^{-6}$ are taken from [8].

3. Simulations

For modelling, a dynamic nonlinear complex system has been selected, consisting of the serial connection of two dynamic nonlinear simple plants described by the second-order nonlinear difference equations:

$$y^{(1)}(k+1) = f_1\left(y^{(1)}(k), y^{(1)}(k-1), y^{(2)}(k-2), u(k), u(k-1)\right), \tag{18}$$

$$y^{(2)}(k+2) = f_2\left(y^{(2)}(k+1), y^{(2)}(k), y^{(2)}(k-1), y^{(1)}(k+1), y^{(1)}(k)\right), \tag{19}$$

where the nonlinear functions f_1 and f_2 are given by the formulae:

$$f_1(x_1, x_2, x_3, x_4, x_5) = \frac{x_1 x_2 x_3 x_5 (x_3 - 1) + x_4}{1 + x_2^2 + x_3^2}, \tag{20}$$

$$f_2(v_1, v_2, v_3, v_4, v_5) = \frac{v_2 v_3 v_5 (v_1 - 1) + v_4}{1 + 2v_1^2}. \tag{21}$$

The system input and the model input of the learning signal is given by:
$$u(k) = \sin(2\pi k / 50) \text{ for } 0 \le k \le 250.$$

The system input and the model input of the testing signal is given by:
$$u(k) = \sin(2\pi k / 250) \text{ for } k \le 250$$

$$u(k) = 0.8\sin(2\pi k / 250) + 0.2\sin(2\pi k / 25) \text{ for } 250 < k \le 500$$

As the global model (GM) and the global model with the quality of local models taken into account (GLM), the 6-layer feedforward neural network (see Fig. 2) of the following structure has been adopted: 5 network inputs; 17 neurons (of the hyperbolic tangent transfer function) in the first hidden layer; 4 neurons

in the second hidden layer; 1 linear neuron in the "additional" (third) layer; 17
and 4 neurons in the fourth and the fifth layers, respectively; 1 linear neuron in
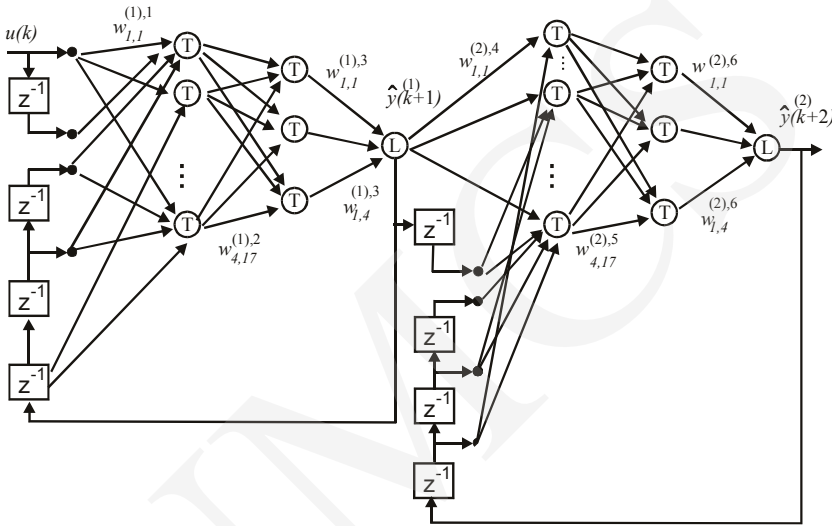the sixth (output) layer of the complex model (shortly, 1(5)-17-4-1(4)-17-4-1).

Fig. 2. The neural network as the (parallel model – working mode)
global model (GM) of the complex system

The following values of the factors of formula (3) in the GM model have
been arbitrarily adopted: $\beta_1 = 0.5$, $\beta_2 = 0.5$. The following values of the factors of
formula (7) in the GLM model have been arbitrarily adopted: $\beta_1 = 0$, $\beta_2 = 1$,
$\alpha_0 = 0.5$, $\alpha_1 = 0.25$, $\alpha_2 = 0.25$.

For learning all the models the modified gradient with the learning rate
$\eta = 0.001$ and the modified *Rprop* algorithm were used. The adjustment of
parameters was performed by computing gradient of the averaged error function
for all 500 discrete time samples.

As an additional criterion of modelling quality assessment, the relative
percentage error RPE (22) has been adopted and calculated for each output of
r-th model:

$$\mathrm{RPE}^{(r)} = \frac{\sum_{k=1}^{K}\left|\hat{y}^{(r)}(k) - y^{(r)}(k)\right|}{\sum_{k=1}^{K}\left|y^{(r)}(k)\right|} \cdot 100\% .\qquad(22)$$

Results of modeling in the form of values of the global criterion Q_g, synthetic
criterion Q_w, the local criteria Q_2, Q_1 and the simple model criterion $Q_g^{(2)}$ (where
$Q_g^{(1)} = Q_1$) after 500 epochs of learning, are shown in Table 1 for both series-

parallel (feedback from the plant) and parallel (feedback from the model) models. The values of RPE after 500 learning epochs are presented in Table 2.

Table 1. Quality criteria after 500 learning epochs

Criterions	Q_g/Q_w	Q_1	Q_2	$Q_g^{(2)}$	Q_g/Q_w	Q_1	Q_2	$Q_g^{(2)}$
Models	Learning – series-parallel model				Testing – parallel model			
GM – gadient alg.	0.509	0.503	-	0.516	0.352	0.243	-	0.470
GLM – gradient alg.	0.428	0.387	0.561	0.382	0.449	0.405	0.519	0.4360
GLM – *Rprop* alg.	0.0147	0.0352	0.0127	0.0054	0.0614	0.144	0.0143	0.0435
GM – *Rprop* alg.	0.0294	0.0366	-	0.0222	0.1710	0.273	-	0.0679
GLM – grad. (2000 epoch)	0.117	0.172	0.124	0.086	0.312	0.404	0.246	0.299

Table 2. RPE after 500 learning epochs for testing data for parallel model

Number of output	$r = 1$	$r = 2$ (locally)	$r = 2$	$R = 1$	$r = 2$ (locally)	$r = 2$
	for gradient algorithm			*for Rprop algorithm*		
Models	[%]	[%]	[%]	[%]	[%]	[%]
GLM	4.83	8.92	7.42	2.79	1.36	2.19
GM	3.55	-	8.16	3.77	-	2.78

The errors in the parallel model for the same epoch number are much smaller by using the *Rprop* algorithm than using the gradient algorithm for the learning.
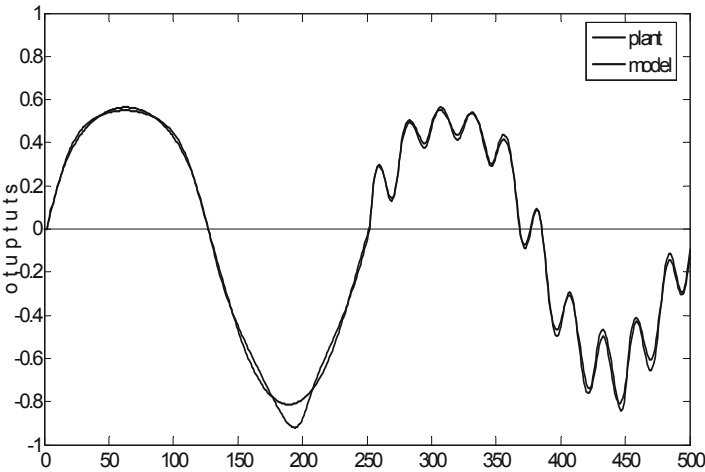

Fig. 3. The outputs of the first local model of GLM and the first simple plant for the testing data

The errors of external outputs of the GM model are bigger than those in the GLM model for both kinds of algorithm. In the GLM model the errors of the output of the second local model ($r = 2$) are lower than those of external output of that model for the *Rprop* algorithm (for gradient after 2000 epoch learning).

Testing signals in the GLM model for gradient algorithm after 500 epoch of learning are shown in Figs. 3 and 4.
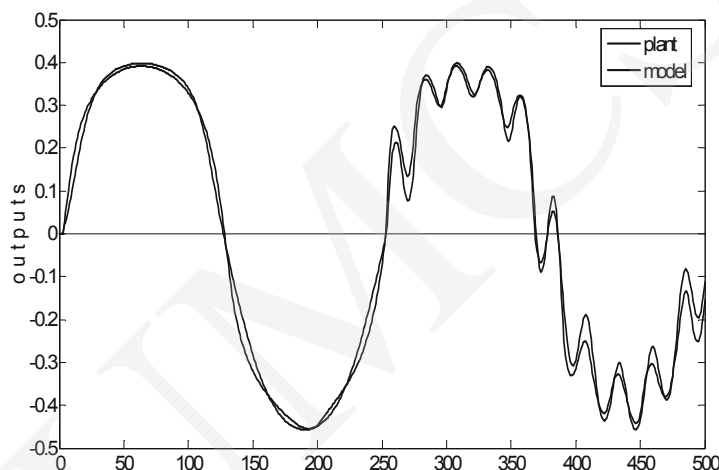

Fig. 4. The external outputs of the GLM model (parallel model) and the complex system for the testing data

Conclusions

In the paper the results of computer simulation for the global model (GM) and the global models with the quality of local models taken into account (GLM) are presented. The quality (accuracy) of the GM model and the GLM model are satisfactorily good. The modified *Rprop* algorithm was much faster than the modified gradient backpropagation algorithm which needed the acceptance of a very small learning rate due to many learning samples. Learning for both discussed algorithm was in reasonable time.

The presented results confirm the possibility of developing global models of dynamic complex systems using the static multilayer feedforward neural networks with TDL.

Further work is to use the proposed method of the neural networks which are the universal approximators and the modified learning algorithms to model a complex system with difficult or unknown dynamic. It is for modelling a number of plants, which can have many inputs and many outputs.

References

[1] Bubnicki Z., *Identification of Control Plants*, Oxford, Amsterdam, N.York: Elsevier, (1980).

[2] Dahleh M.A., Venkatesh Sr., *System Identification for Complex Systems: Problem Formulation and Results*, In Proceedings of the 36th IEEE Conf. on Dec. and Control, (1997).

[3] Józefczyk J., *Decision making problems in complex of operations systems*, Wrocław University of Technology Press, Wroclaw, (2002), in Polish.

[4] Drałus G., Świątek J., *Globally optimal models of complex systems with regard to quality local models.* Proceedings of 14th "International Conference on Systems Science", Wrocław, (2001), 217.

[5] Drałus G., *Modeling of Dynamic Nonlinear Complex Systems Using Neural Networks.* Proceedings of the 15th "International Conference on Systems Science", Wrocław, III, (2004), 87.

[6] Narendra K.S., Parthasarathy K., *Identification and Control of Dynamic Systems Using Neural Network*, IEEE Trans. On Neural Networks, 1(1) (1990) 4.

[7] Riedmiller M., Braun H., *RPROP – a fast adaptive learning algorithm*, Technical Report, University Karlsruhe, (1992).

[8] Zell A., *SNNS – Stuttgart Neural Network Simulator*, *User Manual*, Stuttgart, (1993).