Int. J. Appl. Math. and Comp. Sci., *1999*, *Vol.9*, *No.3*, *599–617*

NEURAL NETWORK EVALUATION OF MODEL-BASED RESIDUALS IN FAULT DETECTION OF TIME DELAY SYSTEMS

PAVEL ZÍTEK*, RENATA MÁNKOVÁ*, JAROSLAV HLAVA*

Model-based fault detection becomes rather questionable if a supervised plant belongs to the class of systems with distributed parameters and significant delays. Two methods of fault detection have been developed for this class of plants, namely a method of functional (anisochronic) state observer and a modified internal model control scheme adopted for that purpose. Both these model schemes are employed to generate residuals, i.e. differences suitable to watch whether a malfunction of the control operation has occurred. Continuous evaluation of residuals is provided by means of a dynamic application of artificial neural networks (ANNs). This evaluation is carried out on the basis of prediction of time series evolution, where the accordance obtained between the prediction and measured outputs is used as a classification criterion. Implementation of both the methods is demonstrated on a laboratory-scale heat transfer set-up, making use of the Real-Time Matlab software.

Keywords: model-based fault detection, anisochronic model, state observer, internal model control, artificial neural networks.

1. Introduction

Model-based fault detection methods have been developed in the last decade making use of two main approaches (Patton *et al.*, 1989), namely state estimation by means of state observers and parameter estimation by applying an identification procedure. Both these approaches evaluate certain differences between a 'faultless model' and the observed real behaviour of the system to be monitored and therefore they need to be provided with models which are sufficiently accurate and true in describing the faultless behaviour of the process watched. As far as the model-process accordance is concerned, the modelling methods usually used coped well with systems with clearly lumped parameters. On the contrary, the model-based approach involves crucial troubles if systems with distributed parameters and significant delays are treated. Using the standard state space, a sufficiently accurate model introduces a rather large number of artificial state variables resulting, among other things, in a complicated state observer scheme.

* Czech Technical University Prague, Department of Instrumentation and Control Engineering, Technická Str. 4, 166 07 Praha 6, Czech Republic, e-mail: zitek@fsid.cvut.cz.

Model-based approaches to fault detection of distributed and time delay systems may become more favourable and serviceable if the conventional state space framework is replaced by a functional extension called anisochronic one (Zítek, 1998b). The essence of this extension consists in considering a segment of state variables as the system state instead of the usual instantaneous state vector. Due to this concept, it is possible to reduce substantially the number of state variables, while a minority of them have to be artificially defined. As a rule, measurable system outputs usually constitute a major part of the state variable vector if this kind of model is applied. A linear form of the anisochronic state space model is as follows:

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \int_0^T \mathrm{d}\boldsymbol{A}(\tau)\boldsymbol{x}(t-\tau) + \int_0^T \mathrm{d}\boldsymbol{B}(\tau)\boldsymbol{u}(t-\tau), \quad \boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} \tag{1}$$

where $\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{y}$ are state, input and output vectors, respectively, τ and T are the delay variable and its upper bound, and $\boldsymbol{A}(\tau)$, $\boldsymbol{B}(\tau)$ are functional matrices the elements of which determine not only the gain coefficients but also delay distribution functions in appropriate relationships. As a rule, these distribution functions are discontinuous, i.e. they are composed of weighted steps. While the conventional state space description represents only a structure of point-concentrated accumulations, the model (1) results in a network of both accumulations (simultaneous integrators) and delay relations (delayors). A particular form of $\boldsymbol{A}(\tau)$ and $\boldsymbol{B}(\tau)$ can be widely diversified, however the Laplace transform of (1) is always available. As far as the transfer properties are concerned, i.e. in the case of zero initial conditions of the functional differential equation in (1), the transform is as follows:

$$s\boldsymbol{x}(s) = \boldsymbol{A}(s)\boldsymbol{x}(s) + \boldsymbol{B}(s)\boldsymbol{u}(s) \tag{2}$$

where

$$\boldsymbol{A}(s) = \int_0^T \exp(-s\tau)\mathrm{d}\boldsymbol{A}(s), \quad \boldsymbol{B}(s) = \int_0^T \exp(-s\tau)\mathrm{d}\boldsymbol{B}(s) \tag{3}$$

On the basis of the above model, most of conventional methods of control system design can be modified and applied to the plants with distributed parameters and delays. Since the functional state space of $\boldsymbol{x}$-segments over $\langle t-T, T\rangle$ keeps the essential properties (separability, consistence) claimed for a state space in general, controllability and observability concepts can be defined for (1) too. Also various ideas applied in the conventional control system design can be modified for the systems with distributed parameters and delays as soon as they are described as anisochronic systems (Zítek and Hlava, 1998). For instance, the methods of inverse-based or internal model control system design can be modified in this way (Zítek and Hlava, 1998). Let us note that the order of (1) is substantially reduced: even quite involved process dynamics can be described by (1) of second order only.

2. Anisochronic State Observer Design

State observers in fault detection serve as parallel models generating residuals, i.e. process-model output differences able to indicate system faults by increasing deviations exceeding some bounds. The convergence of both the observer and process outputs is provided by an available output feedback. If this feedback is properly designed and if an observability condition is satisfied, the state vector is well estimated in real time. Anisochronic state variables in (1) can be selected much more as available outputs, but even so some artificial ones may also appear. Similarly to the conventional state observer structure, its anisochronic modification is designed as follows:

$$\frac{\mathrm{d}\hat{\boldsymbol{x}}(t)}{\mathrm{d}t} = \int_0^T \mathrm{d}\boldsymbol{A}(\tau)\hat{\boldsymbol{x}}(t-\tau) + \int_0^T \mathrm{d}\boldsymbol{B}(\tau)\boldsymbol{u}(t-\tau) + \int_0^T \mathrm{d}\boldsymbol{L}(\tau)\Big[\boldsymbol{y}(t-\tau) - \boldsymbol{C}\hat{\boldsymbol{x}}(t-\tau)\Big] \tag{4}$$

where $\boldsymbol{L}(\tau)$ is an observer feedback matrix, functional again in general. Its dimensions (n, l) are given by the dimensions of the state and output vectors, respectively. Since most of the state variables coincide with the outputs, the output matrix $\boldsymbol{C}$ is usually simple, composed of units and zeros.

Observability condition. The requirement of observability, i.e. that system state variables can really be reconstructed from the measured outputs, is a crucial condition for any observer design. For the system (1) the spectral observability concept (Lee and Olbrot, 1981) can be adopted. System (1) is spectrally observable if and only if the matrices $\boldsymbol{A}(\tau)$ and $\boldsymbol{C}$ satisfy the requirement

$$\operatorname{rank}\begin{bmatrix} s\mathbf{1} - \boldsymbol{A}(s) \\ \boldsymbol{C} \end{bmatrix} = n \tag{5}$$

for any complex s. If this is satisfied, any of the $\boldsymbol{A}(s)$ eigenvalues can be observed from the output measurements by means of the observer (4) (Zítek, 1998b).

Local observer separability. One of the problems encountered in the observer design is a usually rather high number of observer parameters, i.e. the $\boldsymbol{L}(s)$ elements. This number is given by the product nl (the numbers of state and output variables) and with respect to n some of the $\boldsymbol{L}(s)$ elements are redundant as a rule. On the other hand, complex technical systems are composed of co-operating simpler components, which can be viewed as separated units as regards the state observation. The key condition of their separability is that the interrelations between this unit and the other ones are carried out by means of available outputs only, not by the artificially defined state variables. Suppose that the state vector is partitioned as $\boldsymbol{x} = [\boldsymbol{x}_s^T, \boldsymbol{x}_m^T, \boldsymbol{x}_e^T]^T$, where $\boldsymbol{x}_s$, $\boldsymbol{x}_m$, $\boldsymbol{x}_e$ are the inner state variables of the unit 'S', the outer available (measurable) state variables and the other artificial state variables, respectively. The unit 'S' can be observed separately if and only if the adequately partitioned matrix

$$\boldsymbol{A}(s) = \begin{bmatrix} \boldsymbol{A}_s(s) & \boldsymbol{A}_m(s) & \boldsymbol{A}_e(s) \\ & \text{other parts} & \end{bmatrix} \tag{6}$$

has only zeros in $\boldsymbol{A}_e(s)$, i.e. if $\boldsymbol{A}_e(s) = \mathbf{0}$.

If (6) holds, this property means that the unit 'S' is influenced by other system units in a way which always can be expressed by measurable state variables and hence need not be estimated by the observer. It should be emphasized that the anisochronic state formulation of the process model is particularly favourable in selecting the state variables as measurable quantities. That is why this class of models is well suitable for state observer design. An example of anisochronic observer design for a thermal laboratory system is presented in Subsection 5.2.

3. Internal Model Control of Time Delay Systems and Its Use for Generation of Residuals

Anisochronic state observers described in the previous section are a very appropriate means for generating residuals in time delay systems. However, if a controller is not based on state feedback but it uses output feedback only, the state observer is an additional component and the whole control and fault detection system becomes considerably more complex and more difficult to design and implement. For this reason, it can sometimes be more advantageous to use a structure of the control system in which the necessity to add additional components for residuals generation is minimized. Since the Internal Model Control (IMC) structure (Frank, 1974; García and Morari 1982; Morari and Zafiriou, 1989) contains the system model as an inherent part, it is a suitable candidate for this application. Moreover, IMC has already proved to be an effective method for designing robust control systems and, unlike most other developments in modern control theory, it has been widely accepted also by control engineering practitioners.

However, both the theoretical treatment and applications of IMC described in the literature have been limited to control of systems without delays or simple time delay systems with one delay in control only. In exceptional cases where an attempt to design an IMC controller for a more complex time delay system with state delays was done (see e.g. Roduner and Geering, 1996), time delays in the system were approximated by Padé approximants, which resulted in a complicated controller of a very high order. However, these attempts to use delay free approximations for time delay systems are based on reasons that are much more traditional than rational and they actually cause more problems than benefits.

Since this paper is about fault detection in time delay systems, we will first briefly (see Zítek, 1998a; Zítek and Hlava, 1998 for more details) show in the first part of this section that the IMC concept can be extended to general time delay systems with state and control delays, i.e. systems that cannot be controlled by a Smith predictor and for which only a very awkward and ineffective control schemes such as finite spectrum assignment or LQ control for time delay systems have been proposed in the literature until now.

A basic IMC structure is depicted in Fig. 1 with a solid line. In the case of perfect process model accordance, there is a very transparent relation between controller design and closed-loop behaviour. It can be expressed by the following transfer functions:

$$Y(s) = G_{Mu}(s)R^*(s)W(s) + \Big(1 - G_{Mu}(s)R^*(s)\Big)G_{Md}D(s) \qquad (7)$$

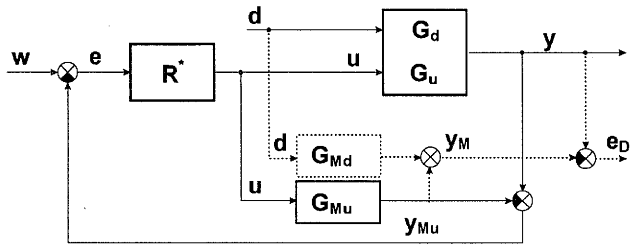

Fig. 1. Augmented IMC scheme.

Lumped delay case. Let us now consider retarded SISO time delay systems with lumped delays in states and in control. This is a subclass of systems described by (4) in which the Stieltjes integral reduces to summation. This class of systems is described by state equations of the form

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{A}_0\boldsymbol{x}(t) + \sum_{i=1}^{l} \boldsymbol{A}_i\boldsymbol{x}(t-\vartheta_i) + \boldsymbol{b}_0 u(t) + \sum_{j=1}^{k} \boldsymbol{b}_j u(t-\tau_j) \\ y(t) &= \boldsymbol{C}\boldsymbol{x}(t) \end{aligned} \tag{8}$$

The corresponding transfer function is given by

$$\begin{aligned} G_{Mu}(s) &= \boldsymbol{C}\left[s\boldsymbol{I} - \boldsymbol{A}_0 - \sum_{i=1}^{l} \boldsymbol{A}_i e^{-s\vartheta_i}\right]^{-1} \left(\boldsymbol{b}_0 + \sum_{j=1}^{k} \boldsymbol{b}_j e^{-s\tau_j}\right) \\ &= \frac{\boldsymbol{C}\mathrm{adj}\left[s\boldsymbol{I} - \boldsymbol{A}_0 - \sum_{i=1}^{l} \boldsymbol{A}_i e^{-s\vartheta_i}\right]\left(\boldsymbol{b}_0 + \sum_{j=1}^{k} \boldsymbol{b}_j e^{-s\tau_j}\right)}{\det\left[s\boldsymbol{I} - \boldsymbol{A}_0 - \sum_{i=1}^{l} \boldsymbol{A}_i e^{-s\vartheta_i}\right]} = \frac{N(s)}{D(s)} \end{aligned} \tag{9}$$

This transfer function is a quotient of two quasi-polynomials and as such it is not a rational function. If the characteristic quasipolynomial $D(s)$ is stable, then since $N(s)$ is a quasi-polynomial with the highest power of s equal to $n-1$ and exponential terms with negative exponents, the transfer function (9) is analytic and bounded in the open RHP, i.e. it is in the function space H^{∞} and it admits an inner-outer factorization

$$G_{Mu}(s) = G_O(s)G_I(s) \tag{10}$$

where $G_O(s)$ is outer and $G_I(s)$ is inner. Both the factors are possibly irrational. In most cases of practical importance, the inner factor reduces merely to a singular function $G_I(s) = e^{-s\theta}$. The transfer function can then be written as

$$G_u(s) = \frac{N_O(s)}{D(s)} e^{-s\theta} = G_O(s)e^{-s\theta} \tag{11}$$

where $D(s)$ and $N_O(s)$ are stable quasi-polynomials. Delays in quasi-polynomials $D(s)$ and $N_O(s)$ are often considerably larger and they may have more significant influence on the dynamic behaviour than the explicit delay θ. This will be also shown in the application example at the end of this paper.

If factorization (11) is possible, an IMC controller can be computed according to the formula

$$R^*(s) = (G_O(s))^{-1} F(s) \tag{12}$$

Robustness issue. Since the controller includes the inverse of infinite dimensional $G_O(s)$, it is also an infinite dimensional system. With $F(s)$ of a sufficiently high order, this controller is proper and it can be shown that it is also causal and the nominal dynamic behaviour of the control system is fully determined by the filter $F(s)$. Owing to the inherent limitation given by the time delay θ, which cannot be overcome by any causal controller, the response can be achieved arbitrarily fast by using a filter with small time constants. However, if filter time constants are small, the performance of this controller becomes very sensitive to modelling errors. Thus, neither nominal stability nor nominal performance are the primary design targets but robust stability and robust performance. An important advantage of IMC lies in the fact that it provides a framework in which robustness can be dealt with explicitly and relatively easily. It can be shown that even in the case of the infinite dimensional system (8) and controller (12) it is possible to use a robust performance condition in the form

$$|S_0(j\omega)w_p(j\omega)| + \left|T_0(j\omega)\bar{l}_m(j\omega)\right| < 1 \quad \forall\, \omega \tag{13}$$

where S_0 is the nominal sensitivity, T_0 stands for the nominal complementary sensitivity $T_0(s) = R^*(s)G_{Mu}(s)$, $S_0(s) = 1 - T_0(s) = 1 - R^*(s)G_{Mu}(s)$, $w_m(j\omega)$ is the performance weight and $\bar{l}_m(j\omega)$ is an upper bound on the multiplicative uncertainty. Using (13) it is possible to find an 'optimal' trade-off between performance and robustness. This is very important because no other method of controller design developed for systems (8) allows for this except for some recently developed and very complicated methods of H^∞ optimal control for time delay systems.

Fault detection supplement. However, the IMC scheme not only allows for an advantageous extension to control complex time delay systems but, since the difference signal which can be used as residual one is inherent in the basic IMC structure, this control structure also offers a suitable starting point for fault detection. Since larger values of error feedback may be caused not only by faults but also by load disturbances, i.e. they may arise even during faultless operation, the basic IMC scheme should be augmented by additional blocks depicted with a dashed line in Fig. 1 and the input to the fault detection block is then given by

$$e_D(s) = y(s) - G_{Mu}(s)u(s) - G_{Md}(s)d(s) \tag{14}$$

Hence the extended internal model can be applied in two roles: as a part of the control system and for computation of the difference e_D. This residual can be employed in a similar manner as in the observer-based methods of fault detection.

In comparison with the state observer technique, this approach offers simplicity and a low cost. Since the disturbance d caused by useful load changes is assumed to be measurable, the control scheme of Fig. 1 could further be augmented by an additional feedforward block compensating for this disturbance. This would bring about a considerable improvement of the disturbance response in the case of perfect process model accordance. However, the performance of the feedforward controller block is rather sensitive to a process model mismatch and, as a result, the use of this feedback/feedforward combination could decrease the robustness and cause a performance deterioration in the presence of errors. Since the controller is required to achieve a reasonable performance even in the case of plant parameter changes and faults, the feedforward block was not implemented.

4. Neural Network Approach to Residual Evaluation

The residuals obtained from the anisochronic observer or from IMC are viewed as symptoms of system faults, i.e. of possible sensor or actuator failures, as well as of process breakdowns. Instead of immediate evaluation, the time series of residuals are subjected to a test by an artificial neural network (ANN) which models the monitored operation on-line. Such a processing takes advantage of watching the cumulative difference between the real and ANN-modelled (predicted) residual signals instead of an alarm caused by exceeding prescribed limits.

The neural detection presented here is based on the *dynamic* performance of the neural detector instead of the more usual *static* pattern recognition approach. The network is set up as a residual *signal predictor* (Adam *et al.*, 1993; Zítek *et al.*, 1999) and the criterion of a 'good' signal prediction becomes an indicator of the process malfunction. The ANN predicts a sequence of residuals over time and the cumulative prediction error, i.e. a sum of squared differences over a selected time interval, indicates discrepancies, if any.

This approach yields several substantial advantages. First, a well-learnt ANN can be used to process available signals of system outputs or control actions and useful loads, and therefore to avoid additional costs for sensors and the measuring technology. Second, a standard detecting means monitors the growing value of a residual signal and gives rise to an accident alarm as soon as a single value exceeds the watched limit. In reality, such a situation need not reflect a real malfunction at all and it can be caused by a single measurement or an information transmission error. The neural predictor overcomes this by reducing the process (plant) dynamics to its constituent part via signal history embodying and thereby it reflects the trespass upon the principle, rather than a random single difference value only. Third, the neural detection is far more sensitive, quicker and it can be tuned based on certain problems in question. Furthermore, it is aimed at signal processing still within 'safe bounds' of differences and it can discover some discrepancies as early as they occur, even if they are not ascertainable by direct observation. Such a competence can bring earlier information on a coming break-down, discover a minute trouble which can be permanently compensated by actuator interventions or which can grow slowly and cause suddenly a plant damage.

Network architecture. The choice of a suitable network architecture has been supported by experiments with various topologies of different size and transfer functions (linear and non-linear feed-forward multilayered perceptrons). Following the previous experience in complex and chaotic signal prediction (Bíla *et al.*, 1997; 1998; Mánková, 1997), quite simple *feed-forward non-linear three-layered networks (perceptrons)* were used. The number of input neurons reflects the history of the residual signal evolution and it is in relation to the estimated order of the process (in the case of model-based simulation), or more generally, to the process complexity. The prediction strategy, i.e. the number of history samples for prediction, creates a set of input nodes and affects significantly the prediction result (the diagnostic ability of the network). In view of the above-mentioned laboratory plant (i.e. a low-order non-linear time delay system) the *one-step-ahead prediction* from a single actual and two previous steps was chosen.

Network learning. Networks were trained by a simple *back-propagation algorithm* with fixed learning rate (due to comparability of results) and without momentum. In the experiments, exposition of time series of 500 samples which describe time evolution over ca 8 minutes (2000 training epochs) was carried out. The networks were trained by time series of characteristic signals belonging to specific operations of the plant. Some attention should be paid to the time series length and to suitable sampling. The series provided a representative training/testing set including all characteristics of the considered events. In the case of a normal operation, they covered the whole width of the 'safe band' of the difference signal and in the case of an unusual event or singularity they illustrated its typical size and dynamics.

The data obtained are very problem-dependent and they vary not only according to certain simulated faults, but also according to plant operating points and useful load changes. For this reason, there is no general rule for finding an optimal network. Every network was tuned 'by hand' by recurrent exposition of certain training data and the experiments confirm the chance to find a suitable solution in most cases.

Detection criterion. The detection is based on a simple classification criterion which reflects the success of a neural network to describe (i.e. to predict) time evolution of the tested signal. Low values of the prediction criterion (a sum of the squared differences between the original and predicted values — SSE) indicate that the signal belongs to the class which the network was learnt on, whereas high values exclude it from that class and identify it as 'abnormal'.

The approach is based on a pure predictive function with saturation by the original output signal (Bíla *et al.*, 1997), i.e. the prediction in every step is computed from actual and previously measured values of the output signal (residuals) and does not have to be supported by information on the plant inputs or internal control interventions.

5. Application Example

5.1. Description of the Controlled System

Practical applicability of the approaches to control and fault detection of time delay systems that were presented in the preceding parts of this paper was tested on

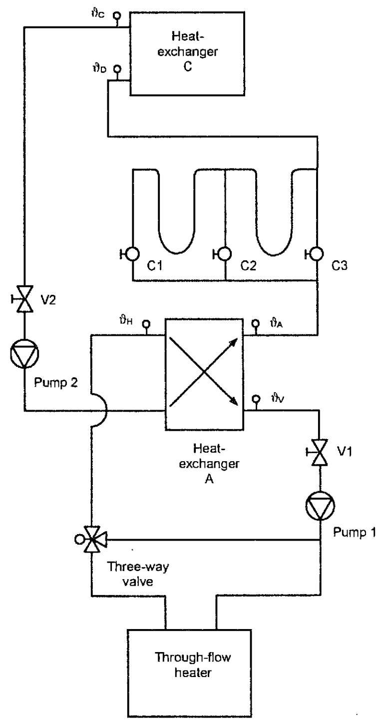

Fig. 2. Scheme of the under consideration.

a *laboratory-scale heating system*. The main components of this system are schematically sketched in Fig. 2. Water in the primary circuit is heated in a through-flow heater. Power consumption of this heater can be changed continuously from zero to a maximum. Hot water from the heater is conveyed to a controllable rotary three-way valve. The position of this valve can be set in a continuous way and it is used as a manipulated variable. This valve mixes the hot water stream with colder water returning from the heat exchanger. Water from the outlet of this valve is conveyed to the countercurrent plate heat exchanger (A) where it heats water in the secondary circuit. Water in the secondary circuit is then cooled in the air-water heat exchanger (C). This air-water heat exchanger is integrated with a fan driven by a DC motor whose speed is controlled by the armature voltage. Thus, besides the electrical power input of the heater and position of the three-way valve, the armature voltage is the third independent input variable. Changes in the fan motor speed, i.e. changes in the intensity of water cooling in the secondary circuit may be interpreted as changes in the useful load of the system.

The heat exchangers are interconnected with long pipes, which results in considerable transportation delays. The delay between the outlet of the heat exchanger A and inlet of the heat exchanger B can be set to three different values depending on the cocks $C1$, $C2$ and $C3$. Delays in both the primary and secondary circuits can further be adjusted more finely by the valves $V1$ and $V2$. Control of this system is accomplished with a PC using the RealTime Toolbox for MATLAB and Simulink. It allows for an easy and flexible programming of an almost arbitrary controller/observer structure including neural and fuzzy ones.

This plant incorporates two heat exchangers connected through long pipes. Thus it includes a number of both distributed and lumped delays. Since connecting pipes are relatively long in comparison with the dimensions of the heat exchangers, transportation delays in the piping have a decisive influence on the overall behaviour of the system. It is possible to consider a model with lumped delays only by approximating heat exchangers by lumped-parameter models (Malleswararao and Chidambaram, 1992; Mozley, 1956). Nevertheless, the mathematical model of the system is relatively complex even with this simplification (it contains several internal state delays) and moreover it is non-linear (Hlava, 1998).

If ϑ_D is a controlled variable, the position of the three-way valve α (its Laplace transform is denoted by $A(s)$) is used as a manipulated variable and the armature voltage of the fan motor u_f is used to change the load, the whole system can be described by the following linearized model:

$$\Theta_D(s) = \frac{N_{11}(s)}{M(s)} e^{-95s} A(s) + \frac{N_{12}(s)}{M(s)} e^{-123s} U_f(s) \tag{15}$$

Clearly, $G_{Mu}(s)$ and $G_{Md}(s)$ in (7) can now be expressed as

$$G_{Mu}(s) = \frac{N_{11}(s)}{M(s)} e^{-95s}, \quad G_{Md}(s) = \frac{N_{12}(s)}{M(s)} e^{-123s} \tag{16}$$

Since this system includes internal delayed feedbacks, $M(s)$ and $N_{12}(s)$ are not polynomials but quasi-polynomials and for the working point used for the experiments described in this paper they are given by the following formulae:

$$\begin{aligned} M(s) &= 80000s^3 + s^2\left(18272 - 3051e^{-17s}\right) + s\left(1103.5 - 468.128e^{-17s}\right. \\ &\quad \left. -6.4e^{-123s}\right) + 15.6 - 8.28e^{-17s} - 1.64e^{-123s} - 0.519e^{-140s} \\ N_{12}(s) &= 217.6s + 55.8 + 19.6e^{-17s} \\ N_{11}(s) &= 50462s^2 + 8598.8s + 154.415 \end{aligned} \tag{17}$$

It is evident that the internal delays in these quasi-polynomials are equally high or even greater than the control delays in series with the transfer functions in (15). Consequently, the dynamics of this system are considerably more complex in comparison with those of delay-free systems or simple systems with time delays in control only.

5.2. An Example of Anisochronic Observer Design

The fault detection methods have been developed primarily for systems with delays and distributed parameters. Their application has been proved on a thermal laboratory set-up with heat exchangers sketched in Fig. 2 and described in Subsection 5.1

in more detail. Circulating water is warmed up to the outlet temperature ϑ_A in the plate heat exchanger A and then cooled down to the outlet temperature ϑ_C in the air-water radiator C. Both the exchangers are rather distant from each other and consequently significant delays arise in their operation. Several anisochronic models have been developed for this system and one of them is used in this section. The dynamics of exchanger A can be described by the second-order anisochronic model:

$$\begin{aligned} T_A \frac{\mathrm{d}\vartheta_1(t)}{\mathrm{d}t} &= k_1 \left[\vartheta_H(t-\tau_H) - \alpha\vartheta_A(t-\tau_M) - \beta\vartheta_C(t-\tau_1-\tau_C)\right] \\ &\quad + \vartheta_C(t-\tau_C) - \vartheta_A(t) \\ T_1 \frac{\mathrm{d}\vartheta_A(t)}{\mathrm{d}t} &= \vartheta_1(t) - \vartheta_A(t) \end{aligned} \tag{18}$$

where ϑ_1 is an inner representative temperature of the warmed up water (an artificial state variable) and ϑ_H is the heating water temperature. The time constants T_A and T_1 express accumulative properties of the exchanger, τ_C is a transport delay between C and A, τ_1, τ_M and τ_H are the delays resulting from the distributed heat exchange parameters, and k_1, α, β are heat exchange static gains. An analogous model structure may be used for the cooler C with the following result:

$$\begin{aligned} T_C \frac{\mathrm{d}\vartheta_2(t)}{\mathrm{d}t} &= B(\vartheta_0)\nu(t-\tau_\nu) - k_2\gamma\vartheta_C(t-\tau_N) \\ &\quad + (1-k_2\varepsilon)\vartheta_A(t-\tau_D) - \vartheta_C(t) \\ T_2 \frac{\mathrm{d}\vartheta_C(t)}{\mathrm{d}t} &= \vartheta_2(t) - \vartheta_C(t) \end{aligned} \tag{19}$$

where the parameter $B(\vartheta_0)$ expresses the influence of the air temperature ϑ_0, k_2 is the heat exchange static gain and the time constants T_C, T_2 have an analogous meaning as T_A, T_1 in (18). The only available outputs of the system are the measured temperatures ϑ_A and ϑ_C, hence the output matrix is

$$\boldsymbol{C} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{20}$$

and the dynamics matrix is as follows:

$$\boldsymbol{A}(s) = \begin{bmatrix} 0 & -\frac{1}{T_A}\left(k_1\alpha e^{-s\tau_M}+1\right) & 0 & \frac{1}{T_A}\left(e^{-s\tau_C} - k_1\beta e^{-s(\tau_C-\tau_1)}\right) \\ \frac{1}{T_1} & -\frac{1}{T_1} & 0 & 0 \\ 0 & \frac{1}{T_C}e^{-s\tau_D}\left(1-k_2\varepsilon\right) & 0 & -\frac{1}{T_C}\left(k_2\gamma e^{-s\tau_N}+1\right) \\ 0 & 0 & \frac{1}{T_2} & -\frac{1}{T_2} \end{bmatrix} \tag{21}$$

It is easy to find out that the observability matrix satisfies the observability condition (5) since the determinant

$$\det \begin{bmatrix} -\frac{1}{T_1} & \frac{1}{T_1}+s & 0 & 0 \\ 0 & 0 & -\frac{1}{T_2} & \frac{1}{T_{21}}+s \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \frac{1}{T_1 T_2} \neq 0 \tag{22}$$

is non-zero for any s. It also is obvious that the matrix $\boldsymbol{A}(s)$ can be decomposed according to (6) for considering the models (18) and (19) as separate units, since none of the artificial state variables ϑ_1 or ϑ_2 represents the mutal influence between the two parts A and C. Therefore it is possible to arrange the state observation system as two local observers for A and C, where the first one can be designed with simple feedback coefficients l_1 and l_2. Its characteristic matrix results in the following simple form:

$$\boldsymbol{A}_{SA}(s) - \boldsymbol{L}_A(s)\boldsymbol{C}_{SA} = \begin{bmatrix} 0 & \frac{1}{T_A}(k_1 \alpha e^{-s\tau_M}+1) - l_1 \\ \frac{1}{T_1} & -\frac{1}{T_1} - l_2 \end{bmatrix} \tag{23}$$

Acceptable observation dynamics can now be achieved by means of prescribing the desired dominant eigenvalues of (23). Let $s_1 = -0.5$ and $s_2 = -0.7$, whence the following feedback coefficients result: $l_1 = 1.034$, $l_2 = 0.914$. With these coefficients the local observation process is aperiodic and fast enough. In a similar way, the other local observer for the part C can be designed. Of course, the cooperation of both the local observers generates a slightly different dynamics from their separated behaviours, but this difference can be kept in quite acceptable limits.

5.3. Internal Model Controller

Since this system is described by irrational transfer functions, the IMC controller computed using (12) will also be irrational. The internal model controller controlling the temperature ϑ_D using the position of the three-way valve as a manipulated variable is given by

$$R^*(s) = \frac{M(s)}{N_{11}(s)} F(s) \tag{24}$$

where a first-order filter is sufficient to make (18) proper:

$$F(s) = \frac{1}{\lambda s + 1} \tag{25}$$

and the controller can then be expressed as

$$R^*(s) = \frac{80000s^3 + 18272s^2 + 1103.5s + 15.6}{(50462s^2 + 8598.8s + 154.415)(\lambda s + 1)}$$

$$- \frac{3051s^2 + 468.128s + 8.28}{(50462s^2 + 8598.8s + 154.415)(\lambda s + 1)} e^{-17s}$$

$$- \frac{6.4s + 1.64}{(50462s^2 + 8598.8s + 154.415)(\lambda s + 1)} e^{-123s}$$

$$- \frac{0.519}{(50462s^2 + 8598.8s + 154.415)(\lambda s + 1)} e^{-140s} \tag{26}$$

or

$$R^*(s) = G_0(s) + G_1(s)e^{-\theta_1 s} + G_2(s)e^{-\theta_2 s} + G_3(s)e^{-\theta_3 s} \tag{27}$$

Its structure is shown in the block diagram of Fig. 3. Besides the controller itself, this diagram also contains blocks of the RealTime Toolbox that allow us to connect it directly to a real system. The structure displayed in Fig. 3 is obviously only a part of the whole control system which includes not only the controller $R^*(s)$, but also the model of the system $G_{Mu}(s)$ and load disturbance model $G_{Md}(s)$ for generating the residuals according to the scheme depicted in Fig. 1. Both $G_{Mu}(s)$ and $G_{Md}(s)$ are given by (16).

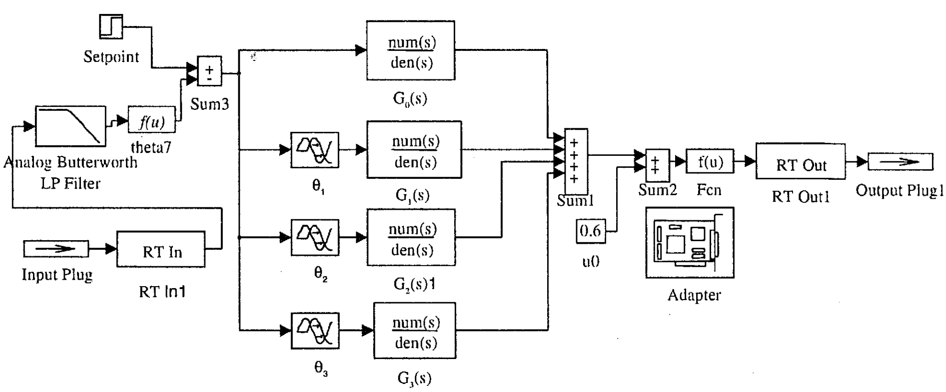

Fig. 3. IMC scheme for the laboratory heating system.

5.4. ANN Evaluation by Neural Predictors

First experiments were aimed at proving the ability of networks to predict a real signal of the above-mentioned nature. The normal process data as well as seven artificial faults were measured simultaneously based on six sensors. After filtering, difference signals were generated and fed to the networks to learn and test (Zítek *et al.*, 1999). As is presumed, a quite simple three-layered feed-forward non-linear network complies best with the non-linear low-order process with delays. A single ANN trained by sets of data series (500 steps/2000 training epochs) is able to distinguish safely the 'normal' difference signals against all exposed 'faulty' signals. The reliable difference in the prediction criterion is of 2–3 orders. The same results are obtainable from all six sensors.

As to the diagnosis, a proper choice of appropriate sensors is the basic preposition of a valid neural verdict. In this case, it was necessary (but also sufficient) to combine data from two representative sensors to cover the whole scope of our fault repertoire. At this stage of research a surprising finding has occurred: in a certain range *'small' faults are better diagnosable* (not detectable!) than the 'big' ones. This is caused by the nature of ANNs; a flexible network trained to 'wild' series of residuals caused by a 'big' fault is able to predict well all the other more 'moderated' residuals of 'small' faults, so evidently distinguishing them is impossible.

Observer generated residuals. In the second experimental step, the idea of *observer generated residuals* processing by ANNs was tested. A local anisochronic state observer was implemented in a real-time linkage to the above-mentioned thermal process. Residuals of the output temperature were processed by the ANN (Zítek *et al.*, 1999). The neural tool was identical to that from the first experiment, i.e. multilayered perceptrons taught by a back-propagation algorithm (intervals of 400 steps/2000 epochs).

The experiments indicate that time series of observer residuals contain important information about the process (plant) dynamics and reflect all its changes well. A repeatedly executed training process enables us to find in every tested case a network which is able to distinguish safely the normal residual signal from any malfunction on the basis of the prediction error. The difference in the values of the criterion is lower than that one in the experiment with simulated residuals, but it is clear enough.

The ability of the diagnosis was tested on two different simulated faults of the same nature on the cooling fan C. The data of each fault served to train the respective network. Both the networks performed their role well: the learnt faulty residuals were predicted with a very low prediction error, the other test signals of strange faults, as well as 'normal' operation residuals, were not well predictable.

IMC residuals. Further, the *residual signal* e_D resulting from the *IMC scheme* can be employed in a similar manner as the one from the observer model. However, its main advantage is the implicity and low costs offered by this approach. In comparison with the state observer technique, no need for artificial state variables' estimation can arise. The better the accordance between the internal model and the real plant properties, the more reliable both control and fault detection is to be expected.

The signals were processed by three-layered non-linear perceptrons again, learnt by exposition of signal time series. In order to test the fault detection ability of the e_D-signal, several plant failures were introduced:

- partial as well as complete fall-outs of the water heater via heating power degradation by 50%, 70% and 100%,

- hydraulic resistance changes in the secondary circuit via appropriate valve adjustments,
- a complete fall-out of the secondary circuit pump,
- a cut-down of the transport delay in the secondary circuit, i.e. setting aside a part of the pipeline via adjustments of appropriate cocks,
- a change in the useful load simulated by a change in the cooling fan speed.

Figure 4 illustrates the 'safe band', i.e. bounds of normal plant operation defined as upper and lower limits of faultless signal evolution. The neural net ALPHA is learnt to identify all signals within these bounds as 'normal.' We can see that the band can be as close as the signal dynamics and the measurement precision allow, i.e. the network is practically an arbitrarily exact detector. According to our wish, ALPHA also classifies the artefacts 'change of operating point' and 'change of useful load' as abnormal (see (a) and (b) in Fig. 5, respectively).

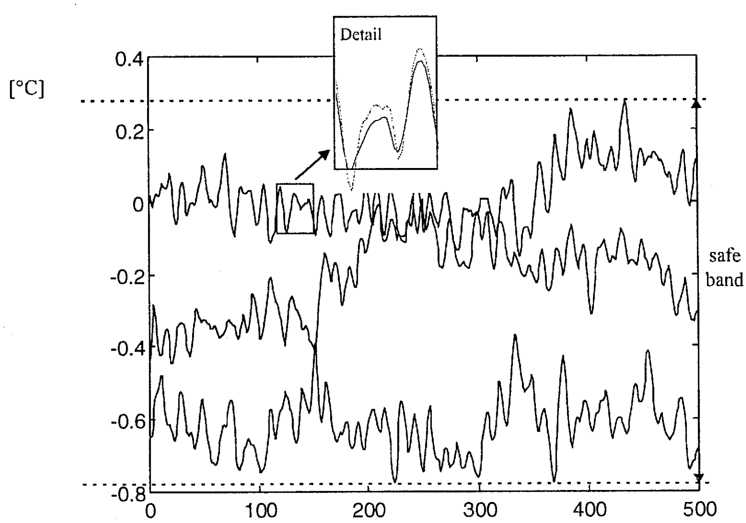

Fig. 4. The class of normal operation. Network ALPHA predicts well signals within certain bounds (SSE< 0.6).

The network BETA follows from ALPHA by further training and can classify the 'change of operating point' to the class of normal operation. Similarly, the network GAMMA understands the 'change of useful load' as a faultless event (Fig. 6).

ALPHA distinguishes real faults based on differences in the value of the prediction criterion SSE (Fig. 7) — the prediction of the normal operation is characterised by average SSE $= 0.326$, while the cut-down of the pipeline and the heater fall-out are indicated by SSE $= 19.006$ and SSE $= 7.726$, respectively.

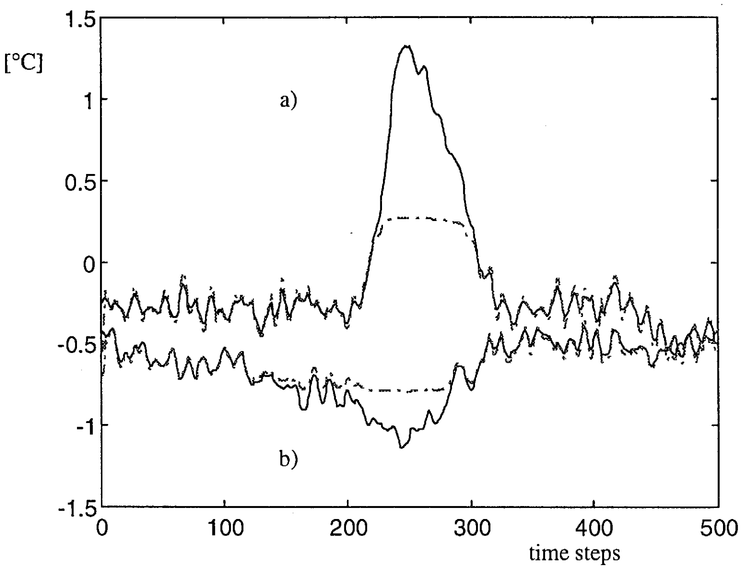

Fig. 5. Distinguishing artefacts from normal operation: ALPHA separates the 'change of operation point' by the value SSE = 37.52 (a) and the 'change of useful load' by the value SSE = 4.21 (b).

6. Conclusions

The fault diagnosis methods presented here have resulted from *integrating* model-based residual generation with ANN prediction and pattern evaluation. Two methods of process-model residual generation have been developed, namely a state observer technique and a modification of the Internal Model Control scheme. In both these cases, applications to controlled processes with distributed parameters and significant delays have been aimed at. To make the application as simple as possible, a recently developed technique of functional state-space models, called anisochronic ones, has been used. This approach helps us to select a low number of state variables, while a major part of them represent measurable process outputs. Thus the model order is reduced and the state estimation is simplified significantly. Of course, the system observability issue has to be modified and viewed in the sense of a spectral observability property. Similarly, the method of IMC-based fault detection uses an anisochronic process model. Thus, since the process transfer properties are followed, transcendental transfer functions serve as a process model and usually define the controller operation as well. A benefit of the IMC approach is that the controller scheme itself serves as a residual generator and any other specialized fault detecting scheme is not needed. Both the models and appropriate residuals have been tested on a laboratory scale thermal system with fairly good results.

A neural predictor applied to signals of dynamic systems can serve as a detector of their normal or faulty operation. In these experiments our efforts were focused on

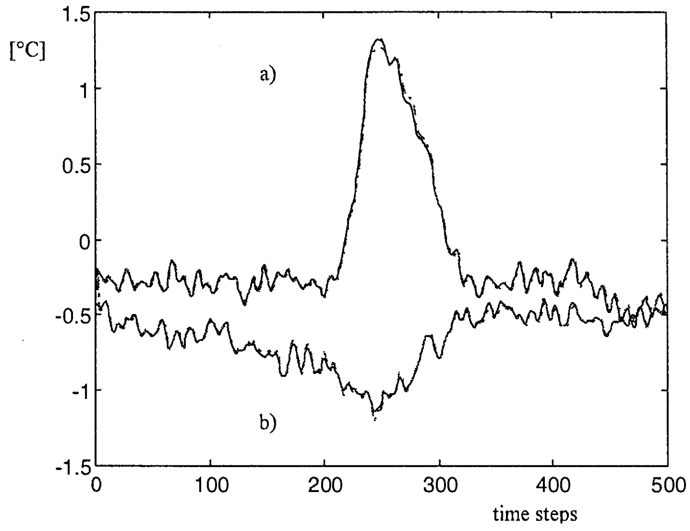

Fig. 6. Training networks: BETA indicates the 'change of operating point' as normal by SSE = 0.36 (a), GAMMA indicates the 'change of useful load' as normal by SSE = 0.24(b).

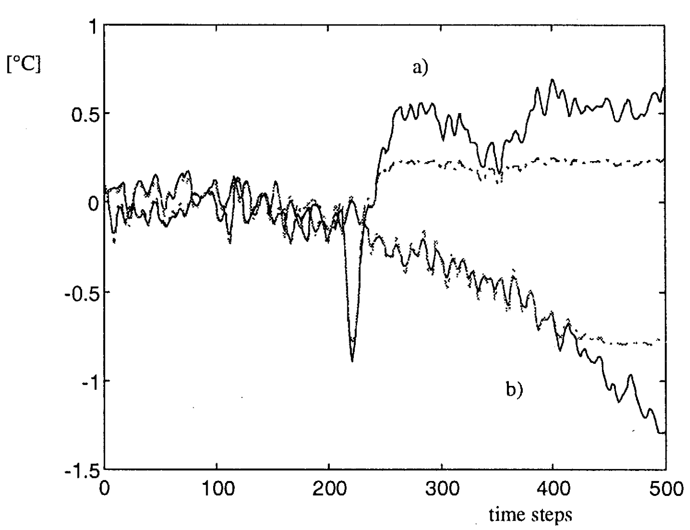

Fig. 7. Classification of real faults. The network ALPHA identifies a 'change of transport delay' caused by excluding pipeline parts (a) and a 'heating fall-out' (b).

neural processing of *observer generated residuals* and *IMC residuals* as characteristic signals of the process. A simple network architecture (a three-layered perceptron), the back-propagation learning and suitable training sets enable us to find a good neural model in every case tested. The ANN taught to predict faultless plant operation detects any deviation from the normal behaviour, even in the case of unclear and visually or statistically similar signals. The experiments with simulated residuals also confirm the possibility to train the ANN on a typical faulty signal and then to distinguish it from the other ones. This result can serve as a basis of a further fault diagnosis and localisation.

The approach seems to be limited to discover faults which demonstrate themselves by crossing some bounds of residual series. That is why it should always be combined with a simple check whether the bounds are crossed. Nevertheless, in comparison with the pattern recognition detection where (even in the case of finding out an abnormal operation only) the risky dominant solution step, i.e. the preliminary feature extraction, is to be overcome, the detection based on neural prediction provides an early, simple and reliable fault announcement without any subjective reinforcement.

With respect to the dynamic character of the ANN application, this idea is suitable for identification of plant (process) malfunctions, rather than those of sensors fall-outs only. Therefore neural detectors can be also implemented in control systems with well-performing or sophisticated controllers that usually disguise by control interventions not only disturbances from outside but also real circuit faults. As shown in Figs. 5 and 6, a neural network can be tuned up to accept an expected event as a normal operation or to classify it as a fault, according to the desired interpretation.

The results support the idea of this dynamic ANN application: the network has proved to be a suitable tool of high sensitivity in fault diagnosis. The integration of model-based generation of encoded information on plant operation and of the dynamic neural prediction in the role of a fault detection tool offers promising results. Once the neural predictor is designed and trained with respect to a selected set of possible faults, it is a simple and low-cost means and it enables the real-time fault detection with an immediate response or even a preventive message.

Acknowledgement

The research was supported by the INCO-COPERNICUS grant (IQ2FD) No. IC15-CT97-0714.

References

Adam O., Zarader J.L. and Milgram M. (1993): *Identification and prediction of non-linear models with reccurent neural network.* — Proc. *IWANN'93*, Sitges, Spain, pp.531–535.

Bíla J., Vitkaj J., Máňková R. and Kváča P. (1997): *Prediction of chaotic signals by means of neural networks.* — Proc. MENDEL 1997 Int. Conf. *Genetic Algorithms, Optimization Problems, Fuzzy Logic, Neural Networks, Rough Sets,* Brno, Czech Republic, pp.228–333.

Bíla J., Kittnar O., Mánková R., Lechmanová M. and Kucha P. (1998): *Modelling of cardio-signals by means of neural networks.* — Proc. Sci. Symp. *Computers in Medicine*, Prague, Czech Rep., pp.23–26.

Frank P.M. (1974): *Entwurf von Regelkreisen mit vorgeschriebenem Verhalten.* — Karlsruhe: G. Braun.

García C.E. and Morari M. (1982): *Internal model control: 1. A unifying review and some new results.* — Industr. Eng. Chemistry Proc. Design and Development, Vol.21, No.2, pp.308–323.

Hlava J. (1998): *Anisochronic Internal Model Control of Time Delay Systems.* — Ph.D. Thesis, Czech Technical University of Prague, Faculty of Mechanical Engineering.

Lee E.B. and Olbrot A. (1981): *Observability and related structural results for linear hereditary systems.* — Int. J. Contr., Vol.34, pp.1061–1078.

Malleswararao Y.S.N. and Chidambaram M. (1992): *Nonlinear controllers for a heat exchanger.* — J. Proc. Contr., Vol.2, No.17.

Morari M. and Zafiriou E. (1989): *Robust Process Control.* — Englewood Cliffs: Prentice-Hall.

Mozley J.M. (1956): *Predicting Dynamics of Concentric Pipe Heat Exchangers.* — Ind. Eng. Chem., Vol.48, No.6, pp.1035–1041.

Mánková R. (1997): *Prediction of Chaotic Signals by Artificial Neural Networks in Application to Analysis of Cardio-signals.* — Ph.D. Thesis, Czech Technical University of Prague, Department of Automatic Control, (in Czech).

Patton R., Frank P.M. and Clark R. (1989): *Fault Diagnosis in Dynamic Systems. Theory and Applications.* — New York: Prentice-Hall.

Roduner Ch. and Geering H.P. (1996): *Modellbasierte Mehrgrössen-Regelung eines Ottomotors unter Berücksichtigung der Totzeiten.* — Automatisierungstechnik, Vol.44, No.7, pp.314–321.

Zítek P. (1998a): *Time Delay Control System Design Using Functional State Models.* — CTU Reports, Vol.2, No.1/98, CTU Prague.

Zítek P. (1998b): *Anisochronic state observers for hereditary systems.* — Int. J. Contr., Vol.71, No.4, pp.581–599.

Zítek P. and Hlava J. (1998): *Anisochronic inverse-based control of time delay systems.* — Proc. Int. Conf. *CONTROL'98*, Swansea, UK, pp.1409–1414.

Zítek P., Bíla J., Šulc B., Mánková R. and Hlava J. (1999): *Neural network processing of observer generated residuals in fault diagnosis of time delay systems.* — IFAC World Congress, Beijing, China, pp.79–84.

Received: 15 December 199
Revised: 23 April 1999