An Indirect Adaptive Neural Control of a Wastewater Treatment Bioprocess via Marquardt Learning

Carlos R. Mariaca-Gaspar and Ieroham S. Baruch

Dept. of Automatic Control, CINVESTAV-IPN, Ave. IPN No 2508, A.P. 14-470, 07360 Mexico D.F., MEXICO
{cmariaca, baruch}@ctrl.cinvestav.mx

(Paper received on July 10, 2007, accepted on September 1, 2007)

Abstract. The paper propose a new Recurrent Neural Network (RNN) model for systems identification and states estimation of a highly nonlinear wastewater treatment bioprocess using the recursive Levenberg-Marquardt learning algorithm. The estimated states of the RNN model are used for an indirect adaptive trajectory tracking sliding mode control. The applicability of the proposed control scheme is applied for continuous wastewater treatment bioprocess model, taken from the literature, where a good convergence and a low Mean Squared Error of reference tracking is achieved.

1 Introduction

The rapid growth of available computational resources led to the development of a wide number of Neural Networks (NN)-based modelling, identification, prediction and control applications, [1], [2]. The main network property namely the ability to approximate complex non-linear relationships without prior knowledge of the model structure makes them a very attractive alternative to the classical modelling and control techniques. The neural networks and the neuro-fuzzy based techniques were successfully applied in several engineering areas as: direct model reference adaptive control of MIMO nonlinear processes, [3]; modeling and control of wastewater treatment process, [4]; neuro-fuzzy control of robotic exoskeleton, [5]. The proposed in the literature neural control gives a good approximation of the nonlinear plants dynamics, better with respect to the other methods of control, but the applied static NNs have a great complexity, and the plant order has to be known. The application of Recurrent NNs (RNN) could avoid these problems and could reduce significantly the size of the applied NNs.

In some early papers, [6], [7] the state-space approach is applied to design a RNN in a universal way, defining a Jordan canonical two or three layer RNN model, named Recurrent Trainable Neural Network (RTNN) and the Backpropagation (BP) algorithm of its learning. Then this general RTNN approach is extended using the Levenberg-Marquardt (L-M) algorithm of learning, [8], [9], applied for direct neural control of mechanical and biotechnological plants. In the present paper we go ahead applying an indirect adaptive sliding mode control of the same biotechnological plant, using RTNN identifier learned by the L-M learning algorithm, executed in real-time, [9].

2 RTNN Topology and Levenberg-Marquardt learning algorithm

A Recurrent Trainable Neural Network model and the dynamic BP learning algorithm, together with the explanatory figures and stability proofs, are given in [7]. The RTNN topology (see Fig. 1), given in vector-matrix form is described by the following equations:

$$X(k+1) = JX(k) + BU(k); \quad J = block - diag(J_i); |J_i| < 1 \tag{1}$$

$$Z(k) = \varphi[X(k)] \tag{2}$$

$$Y(k) = \varphi[CZ(k)] \tag{3}$$

where: Y, X, and U are, respectively, output, state and input vectors with dimensions l, n, m; J is a (nxn)- state block-diagonal weight matrix; J_i is an i-th diagonal block of J with (1x1) dimension. The inequality in equation (1) represents the local stability conditions, [7], [8], imposed on all blocks of J; B and C are (nxm) and (lxn)- input and output weight matrices; $\varphi[.]$ is vector-valued sigmoid or hyperbolic tangent-activation function; k is a discrete-time variable. The stability of the RTNN model is assured by the activation functions and by the local stability condition (1).

The recursive L-M algorithm [8]-[12] is derived by incorporating a regularization term to the recursive prediction error algorithm which becomes:

$$R(k) = \alpha(k)R(k-1) + (1-\alpha(k))(\nabla Y[W(k)]\nabla Y^T[W(k)] + \rho I_{N_w}) \tag{4}$$

Unfortunately, the inversion matrix lemma is now no longer practical and one partial but effective solution is to add a small constant ρ to one of the diagonal elements of $\nabla Y[W(k)]\nabla Y^T[W(k)]$ at time as proposed in [11], [12]. The equation (4) can then be expressed as:

$$R(k) = \alpha(k)R(k-1) + (1-\alpha(k))(\nabla Y[W(k)]\nabla Y^T[w(k)] + \rho Z_{N_w}) \tag{5}$$

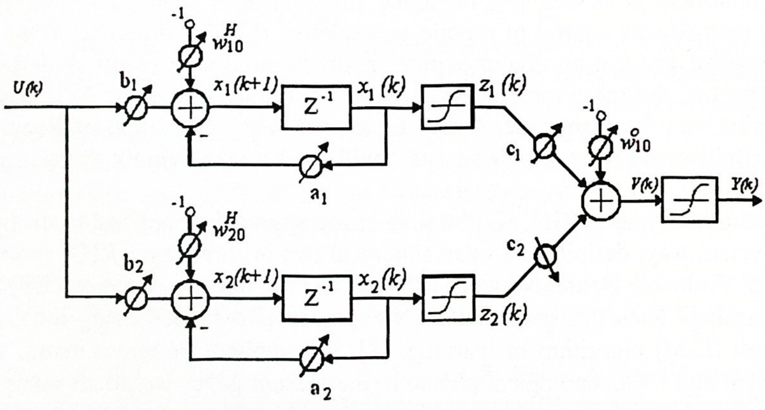

Fig. 1. Recurrent neural network topology (1,2,1).

where Z_{N_w} is a $N_w \times N_w$ diagonal matrix with one non-zero diagonal element which changed from iteration to iteration as follows:

$$z_{ii} = 1, \text{ when} : i = k \bmod(N_w) + 1, \text{ and} : k > N_w \tag{6}$$

$$z_{ii} = 0, \text{ otherwise} \tag{7}$$

With this modification the expression (4) can be re-written in a concise form as:

$$R(k) = \alpha(k)R(k-1) + (1-\alpha(k))(\Omega[W(k)]\Lambda^{-1}(k)\Omega^T[W(k)]) \tag{8}$$

where $\Omega[w(k)]$ is a $N_w \times 2$ matrix with the first column corresponding to $\nabla Y[W(k)]$ and the second column consist of a $N_w \times 1$ vector with one element set to 1, in accordance with equations (9) and (10) above, as it is:

$$\Omega^T[W(k)] = \begin{pmatrix} & \nabla Y^T[W(k)] & \\ 0 & \ldots \; 1 \; \ldots & 0 \end{pmatrix}, \text{ and } \Lambda^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \rho \end{pmatrix} \tag{9}$$

The matrix inversion lemma can now be applied to (8) leading to following recursive Levenberg-Marquardt formulation:

$$S[W(k)] = \alpha(k)\Lambda(k) + \Omega^T[W(k)]P(k-1)\Omega[W(k)] \tag{10}$$

$$P(k) = \frac{1}{\alpha(k)}[P(k-1) - P(k-1)\Omega[W(k)]S^{-1}[W(k)]\Omega^T[W(k)]P(k-1)] \tag{11}$$

Typically the choice of α is in the limits: $0.95 < \alpha < 1$.

As the recursive L-M is based on the Newton method of optimization, it does not needs a stability proof. Next the given up topology and learning are applied for identification and sliding mode control of wastewater treatment bioprocess.

3 Indirect Adaptive Sliding Mode Control Systems Design

Based on the state and parameter estimations, performed by the RTNN identifier, we could propose the following indirect adaptive control scheme, depicted in Fig. 2. The block diagram of that control contained a RTNN identifier and a linear Sliding Mode Controller, designed using the estimated weight parameters.

Let us suppose that the studied nonlinear plant possess the following structure:

$$X_p(k+1) = F(X_p(k), U(k)) \tag{12}$$

$$Y_p(k) = G(X_p(k)) \tag{13}$$

where: $X_p(k)$, $Y_p(k)$, $U(k)$ are plant state, output and input vector variables with dimensions Np, L and M, where L=M is supposed; F and G are smooth, odd, bounded nonlinear functions. The linearization of the activation functions of the learned identification RTNN model, which approximates the plant (see equations (1) to (3)), leads to the following linear local plant model:

$$X(k+1) = JX(k) + BU(k) \tag{14}$$

$$Y(k) = CX(k) \tag{15}$$

where L=M, is supposed.

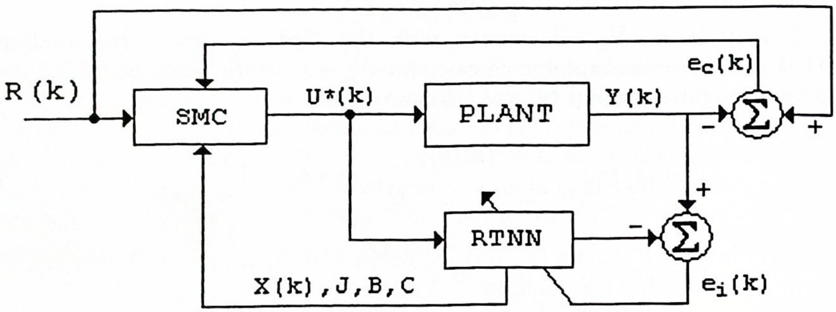

Fig. 2. Block - diagram of the closed-loop system containing neural identifier and sliding mode controller.

Let us define the following sliding surface with respect to the output tracking error:

$$S(k+1) = E(k+1) + \sum_{i=1}^{P} \gamma_i E(k-i+1) \; ; \quad |\gamma_i| < 1 \tag{16}$$

where: S(.) is the sliding surface error function; E(.) is the systems output tracking error; γ_i are parameters of the desired error function; P is the order of the error function. The additional inequality in (16) is a stability condition, required for the sliding surface error function. The tracking error in two consecutive steps is defined as:

$$E(k) = R(k) - Y(k) \; ; \; E(k+1) = R(k+1) - Y(k+1) \tag{17}$$

where R(k) is an L-dimensional reference vector and Y(k) is an output vector with the same dimension.

The objective of the sliding mode control systems design is to find a control action which maintains the systems error on the sliding surface which assure that the output tracking error reaches zero in P steps, where P<N. The control objective is fulfilled if:

$$S(k+1) = 0 \tag{18}$$

Now, let us to iterate (15) and to substitute (14) in it so to obtain the input/output local plant model, which yields:

$$Y(k+1) = CX(k+1) = C[JX(k) + BU(k)] \tag{19}$$

From (16), (17), and (18), it is easy to obtain:

$$R(k+1) - Y(k+1) + \sum_{i=1}^{P} \gamma_i E(k-i+1) = 0 \tag{20}$$

The substitution of (19) in (20) gives:

$$R(k+1) - CJX(k) - CBU(k) + \sum_{i=1}^{P} \gamma_i E(k-i+1) = 0 \tag{21}$$

As the local approximation plant model (14), (15), is controllable, observable and stable, see [7], the matrix J is diagonal, and L=M, the matrix product (CB) is nonsingular, and the plant states X(k) are smooth non- increasing functions. Now, from (21) it is easy to obtain the equivalent control capable to lead the system to the sliding surface which yields:

$$U_{eq}(k) = (CB)^{-1}\left[-CJX(k) + R(k+1) + \sum_{i=1}^{P} \gamma_i E(k-i+1)\right] \tag{22}$$

Following [13], the SMC avoiding chattering is taken using a saturation function inside a bounded control level Uo, taking into account plant uncertainties. So the SMC takes the form:

$$U(k) = \begin{cases} U_{eq}(k), & \text{if } \left\|U_{eq}(k)\right\| < U_0 \\ -U_0 U_{eq}(k) / \left\|U_{eq}(k)\right\|, & \text{if } \left\|U_{eq}(k)\right\| \geq U_0 \end{cases} \tag{23}$$

The proposed SMC cope with the characteristics of the wide class of plant model reduction neural control with reference model, defined by Narendra, [1], and represents an indirect adaptive neural control, given by Baruch, [7].

4 Description of the biological Wastewater Treatment Bioprocess

Wastewater treatment is performed in an aeration tank, in which the contaminated water is mixed with biomass in suspension (activated sludge), and the biodegradation process is then triggered in the presence of oxygen. The tank is equipped with a surface aeration turbine, which supplies oxygen to the biomass, and additionally changes its suspension into a homogeneous mass. After some period, the biomass mixture and the remaining substrate go to a separating chamber where the biologic flocks (biologic sludge) are separated from the treated effluent. The treated effluent is then led to a host environment. The aim is good settling of the biomass in the settler and high conversion of the entering organic material in the bioreactor (see Fig. 3). The main objective of the control system is to keep the recycle biomass concentration close to the reference signal, and this should be achieved in the presence of disturbances and measurement noise acting on the recycle flow rate. A detailed description of all reactions arising in the bioreactor would lead to a high-order model of differential equations [14]. For the control strategy developed in this work a simplified reduced order model is sufficient, as far as it preserves the structural properties of the process, [15]. The model equations are derived using the mass balance of the bioreactor and the settler.

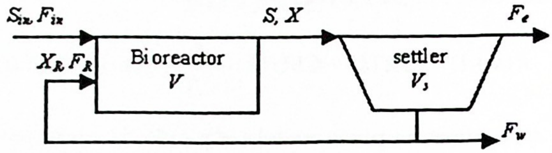

Fig. 3. Biological wastewater treatment with settler.

Mass balance to the bioreactor. The obtained equations are as follows:

$$\dot{X}(t) = \left(\mu(S) - \frac{F_{in}(t) + F_R(t)}{V} - c_d(t) \right) X(t) + \frac{F_R(t)}{V} X_R(t) \tag{24}$$

$$\dot{S}(t) = -\frac{1}{Y}\mu(S)X(t) + \frac{F_{in}(t)}{V} S_{in} - \frac{F_{in}(t) + F_R(t)}{V} S(t), \tag{25}$$

where the state variables are: X(t), biomass concentration; S(t), the substrate measured by the Chemical Oxygen Demand (COD); V is the reactor volume; F_R represents the recycle flow rate (manipulated variable), F_{in} is the influent flow rate; S_{in} is the influent substrate concentration (potential disturbance, also expressed as COD), and $Y>0$ is the yield coefficient. Here $c_d X$ denotes the decay rate of the biomass concentration (which is added in the model to simulate biomass mortality), with $c_d > 0$ as the decay rate parameter. The variable $\mu(\cdot)$ (specific growth rate) is modeled by a Mono-type equation:

$$\mu(S(t)) = \frac{\mu_m(t)S(t)}{K_m(t) + S(t)} \tag{26}$$

where: $\mu_m(\cdot)$ is the maximum growth rate and $K_m(\cdot)$ is the half-saturation constant of biodegradable organic matter. It is the concentration of the substrate for which $\mu = \mu_m/2$. Both parameters are subject to variations.

Mass balance to the settler. It is supposed that none of the biomass is left in the effluent F_e of the settler (see the Fig. 3), so that the whole biomass in the clarifier is settled. The dynamics of the concentration of the biomass in the settler, X_R (t), can be described by the following mass balance equation:

$$\dot{X}_R(t) = \left(\frac{F_{in}(t) + F_R(t)}{V_S} \right) X(t) + \left(\frac{F_W(t) + F_R(t)}{V_S} \right) X_R(t) \tag{27}$$

where: F_W denotes the waste flow rate and V_s is the volume of the settler. We can approximate the settler behavior by:

$$X_R(t) = q(t)X(t) \tag{28}$$

where the parameter q(t) is considered as continuously differentiable and bounded function with bounded inverse, bounded derivative, and q(t) >1 for all t ≥ 0.

Process measurements. The sensor dynamics is modeled by:

$$T_{mm}\dot{X}(t) = -X_m(t) + X_R(t) + n(t) \tag{29}$$

The bioprocess dynamics is corrupted by some white Gaussian noise n(t). The specific model, we consider for the simulations, is obtained after substitution of $X_R(t)$ from equation (28) into equation (24) and has the following form:

$$\dot{X}_R(t) = \left(\frac{\dot{q}(t)}{q(t)} + \mu(t, S(t)) - \frac{F_{in}(t)}{V} - c_d + \frac{q(t)-1}{V}\right)X_R(t) \tag{30}$$

$$\dot{S}(t) = -\frac{1}{Y(t)}\mu(t, S(t))\frac{1}{q(t)}X_R(t) + \frac{F_{in}(t)}{V}S_{in} - \frac{F_{in}(t) + F_R(t)}{V}S(t) \tag{31}$$

Time-varying control reference. The control objective is to assure that the biomass concentration in the recycle flow tracks asymptotically a time-varying reference signal, which is proportional to the influent flow rate and it is assumed to be measurable:

$$X_{Rref}(t) = k_{ref}F_{in}(t) \tag{32}$$

The specific model considered for process simulation is the system of nonlinear differential equations (29), (30), (31), and the Monod-type equation (26), with constant parameters:

$$V = 1{,}5.107\ [\text{l}],\ S_{in} = 300\ [\text{mg COD /l}],\ T_m = 1/12\ [\text{h}] \tag{33}$$

The model uncertainties are taken into account by introducing time-varying parameters as:

$$\mu_m(t) = 0.2 + 0.1\sin(2\pi t/3 + 4\pi/3);\ K_m(t) = 90 + 30\sin(\pi t/2) \tag{34}$$

$$Y(t) = 0.6 + 0.1\sin(\pi t/3 + \pi/3);\ q(t) = 4 + \sin(\pi t/6); \tag{35}$$

$$c_d(t) = 10^{-4}(25 + 5\cdot\sin(\pi t/12)) \tag{36}$$

The control objective is to track the reference signal, given by the equation (32), where the parameters are as follows:

$$k_{ref} = 3.8 * 10^{-3}\ [\text{mgh/l2}];\ F_{in}(t) = 3 * 10^6(1 + 0.25\sin\pi t/12) \tag{37}$$

The initial conditions are always set to:

$$S(0) = 8\ (\text{mgCOD/l}),\quad X_R(0) = 11.4*10^3\ (\text{mg/l}),\quad X_m(0) = 0\ (\text{mg/l}) \tag{38}$$

In order to overcome saturation of the RTNNs, the output and the input of the plant are scaled by the following procedure:

$$y_p = (X_m - 11400)/5700\,;\ F_R = \left[\left(\left[\left(U*7.5\times10^5\right)+3\times10^6\right]F_{Conv}\right)-X_m\right]K_{stab} \tag{39}$$

where the scaling parameters are given by: K_{stab}= 3*10-3 , F_{conv}= 0.0038. These scale factors correspond to the range of the reference signal. Note that the variable U is the bioreactor input control signal, generated by one of the proposed control algorithms, while F_R is the physical process input (the recycle flow rate). Analogous, y_p is the scaled output of the bioreactor, while X_m is the real measured output. Consequently, the reference signal is also normalized following the same procedure (see equations (32) and (37). The reference signal is given by:

$$r(k) = \left(X_{Rref}(k) - 11400\right)/5700 \tag{40}$$

Hence, substituting (37), (44) into (32), and the obtained result in (39), the scaled reference signal is obtained as:

$$r(k) = 0.5\sin\left(\pi k/12\right) \tag{41}$$

The inverse transformation of (46) and the physical bioprocess control signal are as:

$$X_m = 5700 y_p + 11400\,;\ F_R = \left(0.5U - y_p\right)1.71\times10^8 \tag{42}$$

Note that the recycle flow rate F_R is a function of the control variable U, computed by the feedback control with respect to the estimated state and the feedforward control with respect to the scaled plant reference r(k).

5 Simulation Results

A graphical simulation results obtained with the described above wastewater treatment biotechnological plant, obtained using the given up identification and sliding mode control methodology, are shown on Fig. 4. All simulations are performed using the following set of equations: the process description (29), (30), (31); the Mono-type equation (26); the time variable plant parameters (34)–(36); the plant output scaling equation (39); the scaled reference signal equation (41); the scaled plant input equation (42). In all simulations, the severe realistic conditions such as measurement noise are

taken into account by generating a stochastic signal added to the process input and output, both with variance 1200, which is commonly used to simulate noisy measurements.

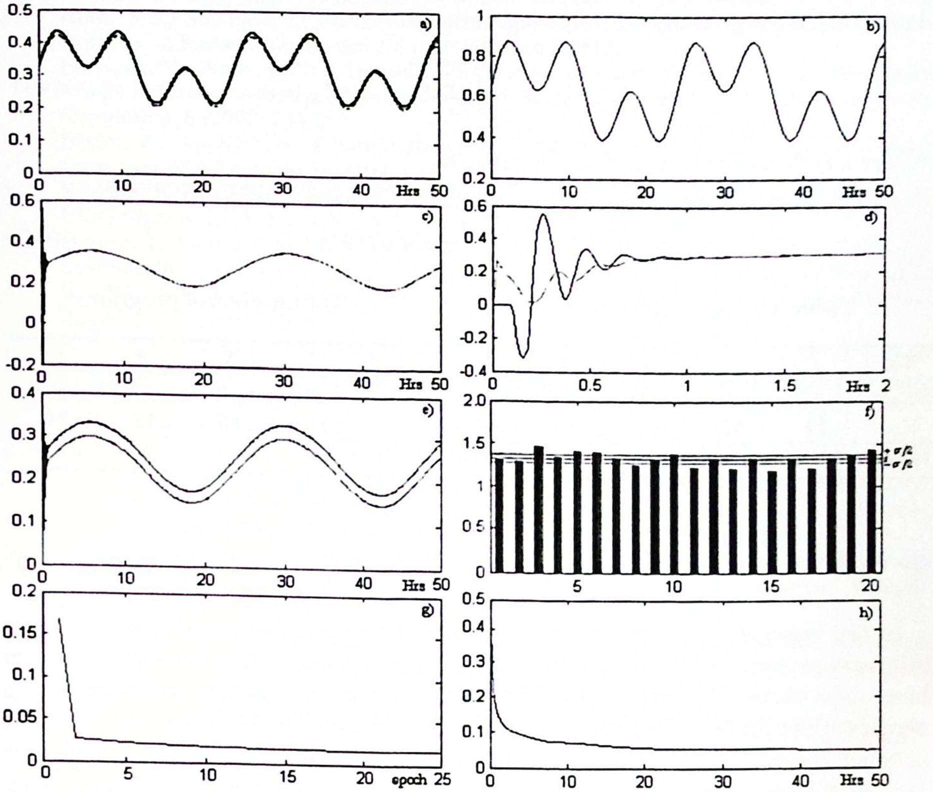

Fig. 4. Graphical results obtained using Indirect SMC and L-M algorithm a) Comparison between the plant output and the reference signal of the control system; b) The control signal; c) System Identification; d) Same as c), but during the first 2 hours. e) States of the plant; f) MSE of control obtained during 20 simulations; g) MSE of identification during 25 epochs; h) The MSE of indirect sliding mode control.

The variance chosen corresponds to 10% noise on the data. The process is simulated over a period of 50 hours, which gives an idea about its periodic behavior (a typical period is about 24 hours) and the period of discretization is set to T_0=0.01h (it is 1 hour of the process time). The parameter used is α=0.95. The activation functions of the hidden and output network layers are hyperbolic tangents. The identification RTNN has topology (1, 2, 1). The results show a good convergence of the system output to the desired trajectory after approximately 1.5 h and a good filtration of the noise which makes a MSE% reduction up to 0.5%. The behavior of the control system in the presence of 10% white Gaussian noise on the plant output could be studied accumulating some statistics of the final MSE% (ξ_{av}) for multiple run of the control program which results are given on Table 1 for 20 runs. The mean average cost for all runs (ε) of control, the standard deviation (σ) with respect to the mean value and the deviation (Δ) are given by the following formulas:

$$\varepsilon = \frac{1}{n}\sum_{k=1}^{n} \xi_{av_k}; \quad \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \Delta_i^2}; \quad \Delta = \xi_{av_k} - \varepsilon \tag{43}$$

Where k is the run number and n is equal to 20.

The mean and standard deviation values of process control are respectively:

$$\varepsilon = 1.3324\ \%; \qquad \sigma = 0.0719\ \% \tag{44}$$

Table 1. Final MSE % of control (ξ_{av}) for 20 runs of the control program.

No.	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
MSE	1.322	1.302	1.468	1.342	1.422	1.402	1.342	1.268	1.314	1.386
No.	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
MSE	1.228	1.342	1.222	1.348	1.214	1.341	1.244	13.42	1.382	1.434

6 Conclusions

In this paper a RTNN model and a dynamic Levenberg-Marquardt learning algorithm are proposed to be applied for identification and state estimation of a nonlinear bioprocess plants. The proposed RTNN model has a Jordan canonical structure, which permits to use the generated vector of estimated states directly for process control. The obtained states are used to design an indirect sliding mode control law. It performs very well under restrictive conditions of periodically acting disturbances, parameter uncertainties and inevitable sensor dynamics. The simulation results, obtained with a continuous wastewater treatment bioprocess plant model, taken from the literature, confirm the applicability of the proposed identification and control methodology.

Acknowledgements

The Ph.D. student Carlos-Roman Mariaca-Gaspar is thankful to CONACYT-MEXICO for the scholarship received during his studies in the Department of Automatic Control, CINVESTAV-IPN, Mexico.

References

1. Narendra, K.S., Parthasarathy, K.: Identification and Control of Dynamic Systems Using Neural Networks. IEEE Transactions on Neural Networks, 1(1) (1990) 4-27.
2. Hunt, K.J., Sbarbaro, D., Zbikowski, R., Gawthrop, P.J.: Neural Network for Control Systems - a Survey. Automatica, 28 (6) (1992) 1083-1112.
3. Frayman, Y., Wang, L.P.: A Dynamically-Constructed Fuzzy Neural Controller for Direct Model Reference Adaptive Control of Multi-Input-Multi-Output Nonlinear Processes. Soft Computing, 6 (2002) 244-253.
4. Boger, Z.: Application of Neural Networks to Water and Wastewater Treatment Plant Operation. ISA Trans. 31 (1) (1992) 25-33.
5. Kiguchi, K., Tanaka, T., Fukuda, T.: Neuro-Fuzzy Control of a Robotic Exoskeleton with EMG Signals. IEEE Trans. Fuzzy Systems 12 (2004) 481-490.
6. Baruch, I., Flores, J. M., Thomas, F., Garrido, R.: Adaptive Neural Control of Nonlinear Systems. In: Dorffner, G., Bischof, H., Hornik, K. (eds.): Artificial Neural Networks-ICANN 2001, Lecture Notes in Computer Science 2130, Springer, Berlin, ISBN 3-540-42486-5 (2001) 930-936.
7. Baruch, I., Flores, J. M., Nava, F., Ramirez, I. R., Nenkova, B.: An Advanced Neural Network Topology and Learning, Applied for Identification and Control of a D.C. Motor. In: Sgurev, V., Jotsov, V. (eds.): Proc. of the First International IEEE Symposium on Intelligent Systems, Varna, Bulgaria (2002) 289-295.
8. Baruch, I., Escalante, S., Mariaca, C. R. RNN Identification and Control of Nonlinear Plants Using the Recursive Marquardt Algorithm. In: J.H.Sossa Azuela, R.Barron Fernandez (eds.) Research in Computing Science, Special Issue: Neural Networks and Associative Memories, ISSN: 1870-4069, IPN, CIC, Mexico City, 21 (2006) 151- 160
9. Baruch, I., Escalante, S., Mariaca, C. R., Barrera, J. Recurrent neural control of wastewater treatment bioprocess via Marquardt learning. In: 10th Computer Applications in Biotechnology, 1 (2007) 285-290
10. Asirvadam, V. S., McLoone, S. F., Irwing, G. W.: Parallel and Separable Recursive Levenberg-Marquardt Training Algorithm. In: Proceedings of the 12th IEEE Workshop on Neural Networks for Signal Processing, (2002) 129-138.
11. Ngia, L. S., Sjöberg J., Viberg, M.: Adaptive Neural Nets Filter Using a Recursive Levenberg-Marquardt Search Direction. IEEE Signals, Systems and Computer, 1 (1998) 697-701.
12. Ngia, L. S., Sjöberg J., Efficient training of neural nets for nonlinear adaptive filtering using a recursive Levenberg-Marquardt algoritm. IEEE Trans. on Signal Processing, 48 (2000) 1915-1927.
13. Young, K.D., Utkin, V.I., and Ozguner, U.: A Control Engineer's Guide to Sliding Mode Control. IEEE Trans. on Control Systems Technology, 7 (3), (1999) 328-342.
14. Sundstrom, D.W., Klei, H.E.: Wastewater Treatment. Prentice-Hall, Englewood Cliffs, NJ (1979).
15. Georgieva, P., Ilchmann, A.: Adaptive λ-Tracking Control of Activated Sludge Processes. Int. J. of Control, 74(12) (2001) 1247-1259.