International Journal of Computers, Communications & Control
Vol. II (2007), No. 2, pp. 121-131

Modelling of Wastewater Treatment Plant for Monitoring and Control Purposes by State – Space Wavelet Networks

Adam Borowa, Mietek A. Brdys, Krzysztof Mazur

Abstract: Most of industrial processes are nonlinear, not stationary, and dynamical with at least few different time scales in their internal dynamics and hardly measured states. A biological wastewater treatment plant falls into this category. The paper considers modelling such processes for monitorning and control purposes by using State - Space Wavelet Neural Networks (SSWN). The modelling method is illustrated based on bioreactors of the wastewater treatment plant. The learning algorithms and basis function (multidimensional wavelets) are also proposed. The simulation results based on real data record are presented.
Keywords: neural network models, model approximation, learning algorithms, waste treatment.

1 Introduction

Biological wastewater treatment plants (WWTP) are very important due to their ability of neutralising results of human activity.

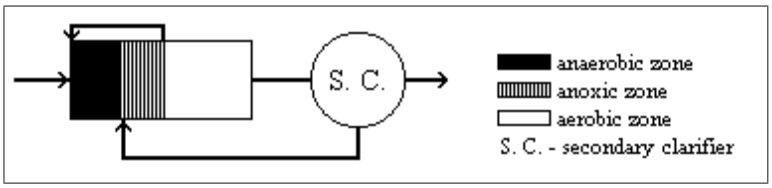

Figure 1: Activated sludge reactor (bioreactor) with secondary clarifier

Typical WWTP consists of three phases of treatment: mechanical, biological and chemical. Example of biological part of WWTP (bioreactor with secondary clarifier) is shown in figure 1. Before biological treatment the wastewater passes through mechanical treatment where coarse particle, inorganic solids and suspended particulate matter are removed. Chemical treatment may be implemented before, after or into biological treatment.

Activated sludge is responsible for nitrogen and phosphorus removing in bioreactor. Biological treatment, due nitrogen and phosphorous removing, consists of three phases: anaerobic, anoxic and aerobic. in figure 1 recirculation (from secondary clarifier to anoxic zone and from anoxic to anaerobic) are shown. Control of that recirculation, the air flow rate to aerobic zone and excessive sludge flow rate is very important for the process quality [8].

Wastewater treatment process is very complex due to its specific features such as: highly non-linear and multiple time scale dynamics, varying influent flow, high dimension of state vector with many states not accessible by hard sensors; see [4] for details.

Due to nationwide regulations, which force high standards on treated wastewater quality, a need for better treatment still exists. There are two solutions of this problem. The first is to enlarge the WWTP and the second, to implement an intelligent control system. ItŠs obvious that the last proposition is better in an economic sense.

The main purpose of control of WWTP is to keep all quality parameters of treated waste under specific norm. Because of complexity of the process, the advanced control technologies are required. The multilevel-multilayer hierarchical control structure was recently derived [1] and further developed in [4].

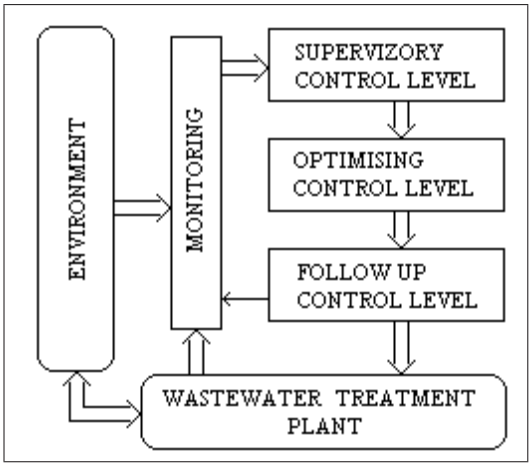

Figure 2: Multilevel-multilayer hierarchical control structure

This structure (figure 2) consists of 3 control levels: supervisory control level, optimising control level and follow up control level; each with different control objectives. Optimising Control Level (OCL) uses a Robust Model Predictive Control (RMPC) algorithm and it is decomposed into three layers each with different control horizon and time scale. This work assembles on modeling the WWTP for RMPC and supervisory control. A good candidate for modelling a dynamical, non-linear system with multiple time scales is a State Space Wavelet Network (SSWN). Using the wavelets as the basis functions solves the multiple time scale problem. The rest of article order goes as follow. Section 2 is a description of problem statement. Wavelet Networks are presented in section 3. A learning algorithm is proposed in section 4. Stability of SSWN is mentioned in section 5. Application and results are shown in section 6 and 7.

2 Problem statement

The paper aim is to verify the possibility of applying the SSWN modelling of a WWTP for monitoring and control purposes.

2.1 Monitoring of WWTP for supervisory control purposes

Monitoring the WWTP state is essential for the supervisory control purposes [4]. Most of the WWTP states are not measurable. Even though the SSWN for state monitoring can be made. First model of WWTP in SIMBA has to be calibrated. Methods for SIMBA model calibration are presented in [2]. When calibrated model in SIMBA is available then it can be used as a data generator for learning the SSWN.

2.2 Progress of the hierarchical intelligent control structure

The layer decomposition of the OCL raises a loss on optimality and certain problems with accommodating the constraints into the layer optimisation tasks. As the new SSWN model has an ability to combine the two time scales (fast and medium), the fast and medium layers can be combined into one layer, hence eliminating the above problems.

3 Wavelet network

Wavelet Network is a neural network with one hidden layer consisting wavelets as the basis function. Wavelets are specific mathematical functions and are described below. Neurons made up of wavelets are called wavelons.

3.1 Wavelets

Any function satisfying the conditions (1) and (2), where $\psi(\omega)$ is a Fourier transform of $\Psi(t)$, is called a mother wavelet. It is required that

$$C_\psi = \int_0^\infty \frac{|\psi(\omega)|^2}{\omega} d\omega < \infty \tag{1}$$

$$\int \Psi(t) dt = 0 \tag{2}$$

Examples of wavelets are Haar wavelet or Morlet wavelet shown in figure 3 and defined by the equations (3) and (4) respectively.

$$\Psi(x) = \begin{cases} 1, & 0 \le x < 0.5 \\ -1, & 0.5 \le x < 1 \\ 0, & \text{elsewere} \end{cases} \tag{3}$$

$$\Psi(x) = \exp(-x^2/2) * \cos(5x) \tag{4}$$

A family of wavelets is made by dilating and translating mother wavelet. See equation (5), where d and

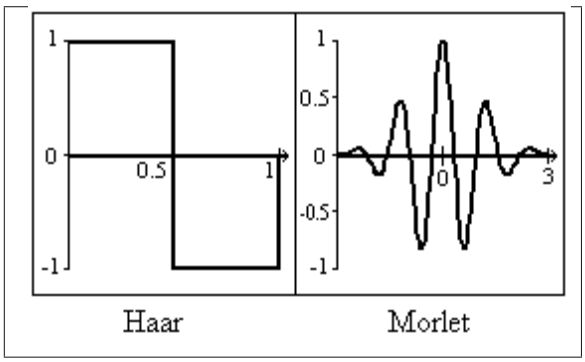

Figure 3: Haar and Morlet mother wavelets.

t represent the dilation and translation parameters respectively.

$$\Psi_{dt}(x) = \sqrt{d}\Psi(d(x-t)) \tag{5}$$

The translation and dilation parameter determine the position (time) and scale (frequency) domain. The wavelets represented by equation (5) may be design as orthonormal.

3.2 Multidimensional wavelets

The multidimensional wavelet (6) was presented first by Zhang and Beneveniste in [11]. In equation (6) $\Psi_S(x)$ is a one–dimensional wavelet, X is a vector and $\Psi(X)$ is a scalar

$$\Psi(X) = \Psi(x_1, ..., x_n) = \Pi_{j=1}^n \Psi_S(x_j) \tag{6}$$

Few months later Zhang [12] presented new multidimensional wavelet - radial wavelet (7) and (8) where $\Psi(X)$ is a one–dimensional wavelet.

$$\Psi(X) = \Psi_S(||x||) \tag{7}$$

$$||x|| = (x^T x)^{\frac{1}{2}} \tag{8}$$

It is logical that for a multivariable function approximation the multidimensional wavelets are desirable. The multidimensional wavelet (6) was tested and some problems were encountered. First, often the optimisation problem to be solved by the learning algorithm was illŰconditioned. Second, a large n relatively easily led to overparameterization.

Only when radial wavelet (equations (7) and (8)) was used the network learning was successfully finalized. The radial multidimensional wavelet, used in the paper, is given by equations (9) - (13) and it is shown in figure 4.

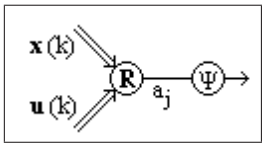

Figure 4: Multidimensional radial wavelon.

$$\mathbf{z}=[\mathbf{x},\mathbf{u}] \tag{9}$$

$$\mathbf{d}_j = [d_{1,j}, ..., d_{K+M+N,j}] \tag{10}$$

$$\mathbf{t}_j = [t_{1,j}, ..., t_{K+M+N,j}] \tag{11}$$

$$\mathbf{A}= diag(\mathbf{d}_j) * (\mathbf{z}-\mathbf{t}_j)^T \tag{12}$$

$$a_j = R(\mathbf{A}(\mathbf{z},\mathbf{d},\mathbf{t})) = (\mathbf{A}^T\mathbf{A})^{\frac{1}{2}} \tag{13}$$

A one Ű dimension wavelet used in this application was a morlet wavelet (4). Proposed structure for a wavelon was described in [12].

3.3 Feed-forward and Input-Output Wavelet Networks.

It was presented in [5] that Feed–Forward Wavelet Network (FFWN) in the form (14) is a universal approximator for any function in Sobolev space.

Wavelet Networks have more freedom then other neural networks because of number of optimised parameters for each wavelon.

$$z = \Sigma_{i=1}^{N} w_i \Psi(d_i(x - t_i)) \tag{14}$$

The wavelet network for one-dimensional input network is described by equation (14), where d_i and t_i are dilation and translation parameters respectively, w_i are linear weights and N is a number of wavelons. A feed forward wavelet network (figure 5) was presented first by Zhang and Beneveniste [11]. A Dynamical Wavelet Network (DWN) was presented in [7]. The DWN structure is shown in figure 6. It is the Input–Output structure, which is commonly used but it has some drawbacks that do not permit this DWN to be used for WWTP modelling. Since 1992 many wavelet networks were presented and used in different places of human activity.

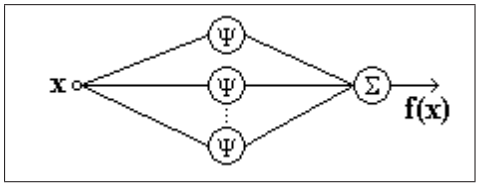

Figure 5: Feed-forward wavelet network.

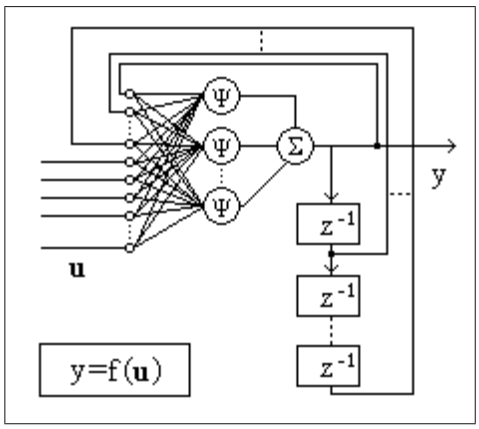

Figure 6: Dynamical wavelet nertwork.

3.4 State Space Wavelet Network

State Space Wavelet Network (SSWN) is not as well known as FFWN or DWN but it has specific advantages over mentioned networks. Due to its of internal state space component the SSWN, better captures the modeled plant structure. Therefore the modeling error is smaller and the learning process is faster. Nevertheless, there is no proof for such a network to be the universal approximator.

SSWN with multidimensional wavelets is shown in figure 7 and described by the equations (15) and (16), where N is a number of outputs, $M+N$ is a number of state variables, K is a number of inputs (control and disturbance), L is a number of wavelons. A number of parameters to be estimated during learning process are then $(2*L*(N+M+K)+L*(N+M))$. It combines the state space architecture of dynamic neural network [10] with the multidimensional wavelons as the processing nodes.

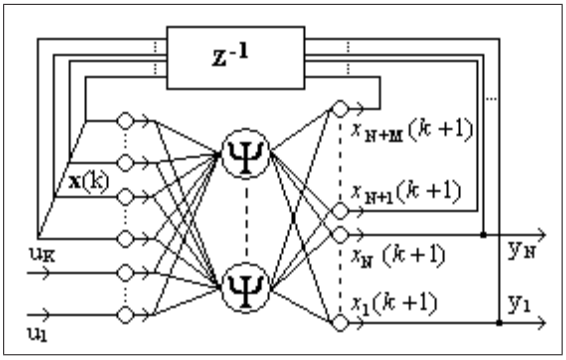

Figure 7: Structure of State Space Wavelet Network with multidimensional wavelons

$$\forall_{i=1}^{N+M} x_i(k+1) = \Sigma_{j=1}^{L} w_{i,j} * \Psi_j(x(k), u(k)) \tag{15}$$

$$\forall_{i=1}^{N} y_i(k) = \Sigma_{j=1}^{L} w_{i,j} * \Psi_j(x(k), u(k)) \tag{16}$$

4 Learning Algorithm

During searching for the right structure of wavelet network, selecting the learning algorithm was also investigated. Searching over a large set of combinatorial optimization algorithms included: Simple Genetic Algorithm (SGA), SGA with elitism, Evolutionary Algorithm (EA), EA with SBX crossover,

Simulated Annealing (SA) and parallel hybrid of SA and EA with the use of computer grid technology. The SA algorithm turned out to be the best for our problems.

4.1 Simulated annealing

Whilst inspiration for GA and EA was in biology (genetic), the SA was inspired by thermodynamics (Statistical Mechanics). The algorithm was motivated by the growing mechanism of a single crystal from a melt [6]. It was found that slow cooling (annealing) of melted metal goes to low state of energy while fast cooling does not.

A simple algorithm based on Monte Carlo search was proposed in (Metropolis 1953) which then became an important part of the SA algorithm [6]. This algorithm works on a chain of atoms (S). In each step one atom from the chain is disturbed and the new chain energy $E(S')$ is calculated. A difference between the chain S and S' energies is calculated by equation (17).

$$\Delta E = E(S') - E(S) \tag{17}$$

If $\Delta E < 0$ then new chain S' is accepted; otherwise the new chain is conditionally accepted with a probability given by Boltzmann probability factor (18), where k_B is a BoltzmannŠs constant and T is a temperature factor. Whilst T is not real temperature and k_B is a constant then product $k_B * T$ may be replaced for practical implementation by single factor T.

$$P(\Delta E) = \exp(-\Delta E/(k_B T)) \tag{18}$$

The annealing schedule was added in [6] in order to formalise the SA algorithm. The SA algorithm works iteratively as follow. For given temperature T_0 the Metropolis Monte Carlo (MMC) method is applied. When chain of atoms is said to be stable then new T is computed and MMC is applied; this procedure goes as long as the temperature reaches 0 or the energy obtains optimum. For the new T the computation-annealing schedule is given by equation (19), where T_0 is initial temperature and k is SA iteration counter. For practical matter chain of atoms is said to be stable when MMC iteration counter obtain an established number.

$$T(k) = \frac{T_0}{1 + \ln(k)} \tag{19}$$

4.2 A simple method for SSWN learning with simulated annealing algorithm

In the paper the learning procedure was implemented as follow:
– Initialize a chain of atoms,
– Set the initial temperature,
– Establish a number of MMC iterations needed to obtain a stable chain,
– Run iterative SA algorithm.

Initialization of chain of atoms is made using a random number generator with upper and lower constrains for weights and wavelet coefficients. The initial temperature is chosen by the user (between 0 and 1). Number of MMC iterations is proportional to length of the chain of atoms. In optimisation only one chain of atom should be used because convergence of SA is independent of initialization due to exploring nature of the method.

Before the SA is applied the energy function must be described. Energy function is given by equation (20), where Y is the plant output and $Y^{\nsim}$ is a SSWN output, N is a number of SSWN outputs, J is a number of samples taken for network learning.

$$E = \frac{1}{N * J} \Sigma_{i=1}^{N} \Sigma_{j=1}^{J} |Y_{i,j} - Y_{i,j}^{\nsim}| \tag{20}$$

It is important to normalize the data before learning the network.

5 Stability of SSWN

It was proved that under certain conditions the State Space Neural Networks (SSNN) can be made stable [9]. The sufficient conditions for the stability suitable constrain the network weights.

It is not a subject of this work to proof the SSWN stability. However, we shall demonstrate by simulation that the stability is in place if the network parameters are constrained. A discharge of a randomly chosen initial state of the network is illustraed in figure 8. After the discharge has been finished the network output accurately follows the plant output.

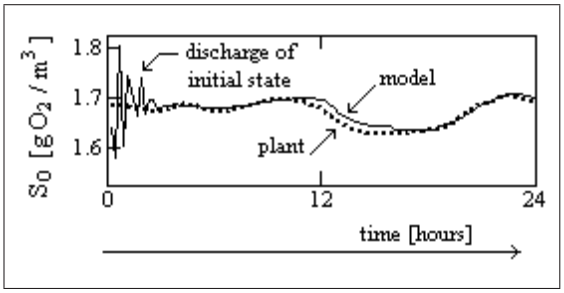

Figure 8: Discharge of randomly chosen initial state of SSWN

6 Application

After the SSWN structure and learning algorithm have been chosen the inputs, outputs, size of the state vector and the number of wavelons must be fixed.

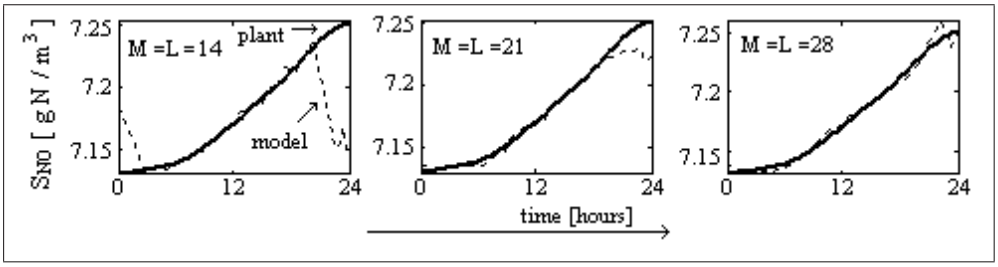

Figure 9: Learning results for different M and L

The presented SSWN was applied to wastewater treatment plant modelling. The modelled plant consists of 4 control inputs (flow rate to aerobic zone, 2 recirculation flow rates and excessive sludge flow rate), 4 measured or estimated disturbance inputs, including inflow and quality (COD, BOD, TSS), and 4 outputs: outflow (Q), concentrations of nitrate and nitrite nitrogen (S_{NO}) and $NH_4^+ + NH_3$ nitrogen (S_{NH}) in effluent and concentration of oxygen in aerobic zone (S_O).

Still the number of state space variables M and number of wavelons L are unknown. Fortunately we know the size of state of the modelled plant, which is 14 for each zone of biological treatment. Therefore M was searched in a set of values: 14, 21 and 28; the larger M implied a huge set of optimized parameters. In order to reduce the computational burden M wavelons were applied. The network was parameterized by 980, 1911 and 3136 parameters, respectively. Results of the learning (limited to single output) for these three parameterization examples are shown in figure 9. Finally M and L were selected as 28 to give small modelling error and acceptable learning time.

7 Results

The results of long term learning are shown in figure 10. The learning time was around 36 hours with simulated annealing algorithm. The bold line shows the modelled plant output while the dashed line illustrates the network output. The mean modelling error was about 1.84

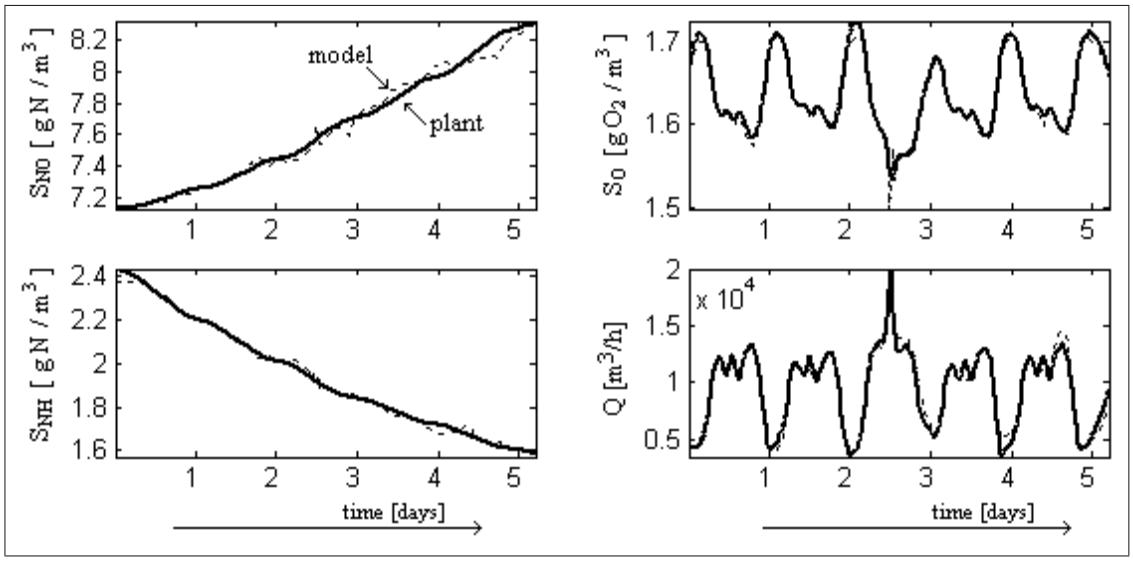

Figure 10: Comparison of model and plant outputs

8 Conclusions

A structure of a dynamical wavelet network called State Space Wavelet Network (SSWN) has been proposed. The Multidimensional Radial Wavelon has been proposed as the network processing nodes. The Simulated Annealing for the SSWN learning has been derived and validated by application to modelling the wastewater treatment plant.

It has been demonstrated that such a network is able to approximate this dynamic, nonlinear, not stationary process with several different time scales.

References

[1] M. A. Brdys, M. Grochowski, K. Duzinkiewicz, W. Chotkowski, Y. Liu "Design of control structure for integrated wastewater treatment plant Ű- sewer systems", *International Conference on Technolog, Automation and Control of Wastewater and Drinking Water Systems TiASWiK'02* Gdansk – Sobieszewo, Poland, June 19-21 2002.

[2] W. Chotkowski, J. Mźkinia, M.A. Brdys, K. Duzinkiewicz, K. Konarczak "Mathematical modelling of the processes in integrated municipal wastewater systems", *Proc. of the 9th IFAC/IFORS/IMACS/IFIP Symposium on Large Scale Systems: Theory and Applications, Bucharest*, July 18-20 2001.

[3] I. Daubechies *Ten Lectures on Wavelets* CBMS-NSF Regional Series in Applied Mathematics, SIAM, Philadelphia, 1992.

[4] M. Grochowski, M.A. Brdys, T. Gmiñski "Intelligent control structure for control of integrated wastewater systems" *IFAC 10th Symposium Large Scale Systems: Theory and Applications* Osaka Ű Japan July 26-28 2004.

[5] A. Juditsky, Q. Zhang, B. Delyon, P-Y. Glorennec, A. Beneveniste "Wavelets in identification" *Rapport de recherche Nř2315* 1994.

[6] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi "Optimization by Simulated Annealing" *Science vol. 220* pp. 671-680, 1983.

[7] Y. Oussar, I. Rivals, L. Personnaz, G. Dreyfus "Training Wavelet Networks for Nonlinear Dynamic Input-Output Modeling" *Neurocomputing, vol. 20*, pp. 173-188, 1998.

[8] G. Olsson, R Newell *Wastewater Treatment Systems. Modelling, Diagnosis and Control.* IWA Publishing, London, 1999.

[9] E.N. Sanchez, J.P. Perez "Input-to-State Stability (ISS) Analysis for Dynamic Neural Networks" *IEEE Transactions On Circuits And Systems Ű I: Fundamental Theory And Applications, vol. 46, No. 11*, pp 1395 Ű 1398, November 1999.

[10] J.M. Zamarreno, V. Pastora "State space neural network. Properties and applications." *Neural Networks, vol. 11*, pp 1099-1112, 1998.

[11] Q. Zhang, A. Beneveniste "Wavelet Networks" *IEEE Trans. on Neural Networks, vol. 3, num. 6*, pp 889-898, Nov. 1992.

[12] Q. Zhang "Wavelet Network: the Radial Structure and an Efficient Initialization Procedure" *Technical Report of Linköping University, LiTH-ISY-I-1423*, October 1992.

[13] J. Zhao, B. Chen, J. Shen "Multidimensional non-orthogonal wavelet basis function neural network for dynamic process fault diagnosis" *Computer and Chemical Engineering vol. 23* pp. 83-92, 1998.

Adam Borowa, Krzysztof Mazur
Gdansk University of Technology,
Department of Automatic Control,
ul. G. Narutowicza 11/12, 80 952 Gdansk, Poland
E-mail: aborowa@ely.pg.gda.pl, kmazur@ely.pg.gda.pl

Mietek A. Brdys
The University of Birmingham, School of Engineering,
Department of Electronic, Electrical and Computer Engineering,
Birmingham B15 2TT, UK
E-mail: m.brdys@bham.ac.uk
Received: March 12, 2007

Adam Borowa received his M.Sc. degree in Control Engineering in 2002 from Electrical and Control Engineering Department at Gdansk University of Technology. Soon after he became a Ph.D. student in this Department. During the period of 2001 to 2002 he served one's apprenticeship on Wastewater Treatment Plant at Swarzewo. He published 6 publications. Mainly he focuses on modelling and monitoring of large scale systems, especially processes with many time scales.

Mietek A. Brdys received the M.Sc. degree in Electronic Engineering and the Ph.D. and the D.Sc. degrees in Control Systems from the Institute of Automatic Control at the Warsaw University of Technology in 1970, 1974 and 1980, respectively. From 1974 to 1983, he held the posts of Assistant Professor and Associate Professor at the Warsaw University of Technology. In 1992 he became Full Professor of Control Systems in Poland. Between 1978 and 1995, he held various visiting faculty positions at the University of Minnesota, City University, De Montfort University and University Polytechnic of Catalunya. Since January 1989, he has held the post of Senior Lecturer in the School of Electronic, Electrical and Computer Engineering at The University of Birmingham. Since February 2001 he has held the post of Full Professor of Control Systems in the Department of Automatic Control at Gdansk University of Technology. He has served as Consultant for Honeywell Systems and Research Center in Minneapolis, GEC Marconi and Water Authorities in UK, France, Spain, Germany and Poland. He is Head of Interdisciplinary Research Network on Decision Support and Control Systems at The University of Birmingham and Head of Intelligent Decision Support and Control System Group at Technical University of Gdansk. His research is supported by the UK and Polish Research Councils, and industry and European Commission. He is author and co-author of about 200 refereed papers and six books. His current research includes intelligent decision support and control of complex uncertain systems, robust monitoring and control, softly switched robustly feasible model predictive control. The applications include environmental systems, technological processes, autonomous intelligent vehicles and defence systems. He is a Chartered Engineer, a Member of the IEE, a Senior Member of IEEE, a Fellow of IMA and a Vice-Chair of IFAC Technical Committee on Large Scale Complex Systems. He the IPC Chair of the 11th IFAC Symposium on Large Scale Complex Systems, Gdansk, July 23-25, 2007.

Krzysztof Mazur received his M.Sc. degree in Control Engineering from Electrical and Control Engineering Department at Gdansk University of Technology in 2005. Currently a Ph.D. student in this Department. His research interests are in the areas of modelling, control and monitoring of large scale systems. Co-author of 4 publications.