IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications
5-7 September 2005, Sofia, Bulgaria

Neuro-Fuzzy Modelling on Experimental Data in Anaerobic Digestion of Organic Waste in Waters

Snejana Yordanova [1], Nelly Noikova [2], Rusanka Petrova [1], Plamen Tzvetkov [1]
1) Faculty of Automation, Technical University of Sofia, 8 Kl.Ohridski blvd., 1000 Sofia, Bulgaria,
e-mail: sty@tu-sofia.bg
2) Rolf Nevanlinna Institute, Department of Mathematics and Statistics, PO Box 68, FIN-00014,
University of Helsinki, Finland, e-mail: nelly.noykova@helsinki.fi

Abstract - The aim of the present paper is to develop neuro-fuzzy models of the anaerobic organic digestion process in wastewater treatment from laboratory and simulated experiments accounting for the variable organic load, ambient influence and microorganisms state in MATLAB environment. The main contributions are determination of significant model parameters via graphical sensitivity analysis, simulation experimentation and design and study in MATLAB of two "black-box" models for the biogas production rate, based respectively on classical feedforward backpropagation and Sugeno fuzzy logic neural networks. The models can find application in process prediction, optimization and control.

Keywords - Anaerobic digestion of organic waste, neuro-fuzzy modelling, sensitivity, simulation

I. Introduction

The anaerobic digestion (methane fermentation) of organic waste is the last stage of water depollution technology, in which organic matter (animal litters, plant sludge, industrial and domestic waste) is mineralised by microorgamisms known as hydrogen acceptors in the absence of oxygen to safely disposable in the environment substances. The complementary product biogas consists mainly of methane and is currently considered as one of the cleanest nonpolluting fuels that can effectively substitute oil or gas in the existing equipment without adjustments. The overall conversion is carried out by a mixture of microorganisms through several biochemical reactions in series and in parallel [1]-[8]. Advantages of the anaerobic digestion are the higher organic loads treated, the smaller amount of sludge produced, the energy recovery via utilization of the biogas, the reduced operating costs as there is no need of oxygen supply and control of its value.

The anaerobic wastewater treatment application is still not very popular due to the process complexity, which makes its mathematical description and hence its control and optimisation difficult. The models used are nonlinear both in terms of parameters and variables, and nonstationary. The parameter identification encounters various problems [2]-[6] due to the specific features of the microorganisms, low reproducibility of the experiments, limited number of time-consuming and expensive measurements and complex laboratory analyses, noisy experimental data, great number of model parameters, etc.

There arises the need for new types of models of the processes, which describe the nonlinear time-varying behaviour combining knowledge of plant experts, measurements and operational experience. Therefore for the purposes of control, prediction and optimisation the fuzzy and neural "black box" model approach to the anaerobic process modelling is encouraged. The fuzzy sets theory offers a methodology for representing heuristic expert knowledge in a computable way by linguistic labels implemented in linguistic rules thus dealing with uncertainties and avoiding complex mathematical relationships [9], [10]. The fuzzy inference process involves membership functions, fuzzy logic operators and knowledge rules. The membership functions (MFs) allow representation of a degree of membership to a fuzzy set, associated to a linguistic label, for a given input numerical value. The rules if-then introduce the expert knowledge in a computable way by means mainly of the operators "and" and "or". The fuzzy set and fuzzy logic theory have been successfully applied to different complex process modelling, prediction and control [11], [12]. The subjectiveness in the choice of the number and the type of MFs as well as their allocation, and in the rule base development can be avoided when input-output experimental data is available and an artificial neural network (ANN) trained [12], making use of the basic advantages of multilayer ANNs with nonlinear activation functions to learn from experimental data, to cluster it, to adapt and generalize in mapping nonlinear relationships.

The aim of the recent investigation is to build neuro-fuzzy models of the anaerobic organic digestion process in wastewater treatment on the basis of laboratory and simulated experiments accounting for the variable organic load and process parameters in MATLAB environment.

II. Problem Formulation

The anaerobic digestion is commonly viewed upon as a three-stage process: hydrolysis and liquefaction of the large insoluble organic molecules; acidogenesis, and methanogenesis [1]-[4]. The process takes place under prescribed temperature and pH since the acidogenic bacteria are sensitive to temperature changes while the methanogenic bacteria cannot tolerate pH fluctuations. In recent years more and more complex mathematical models have been introduced in order to better present the biodegradable processes [1]-[3]. Here the fifth order Barth-Hill nonlinear model with specific growth rates μ_1 and μ_2 of the acidogenic and the methanogenic bacteria respectively in the form of Monod is accepted as an average model to fit the data from a number of laboratory experiments in a continuously stirred tank bioreactor with highly concentrated organic pollutants (cattle wastes) at mesophilic temperature [1], [7], [8] since it reflects the multistage process and the diverse groups of microorganisms involved:

$$\left|\begin{array}{l} \frac{dS_o}{dt} = -DS_o - \beta X_1 S_o + DY_p S_{oi} \\ \frac{dX_1}{dt} = (\mu_1 - k_1 - D)X_1 \\ \frac{dS_1}{dt} = -DS_1 + \beta X_1 S_o - \frac{\mu_1 X_1}{Y_1} \\ \frac{dX_2}{dt} = (\mu_2 - k_2 - D)X_2 \\ \frac{dS_2}{dt} = -DS_2 + Y_b \mu_1 X_1 - \frac{\mu_2 X_2}{Y_2} \\ Q = Y_g \mu_2 X_2 \\ \mu_1 = \frac{\mu_{1\max} S_1}{k_{s1} + S_1}, \mu_2 = \frac{\mu_{2\max} S_2}{k_{s2} + S_2} \end{array}\right. \quad (1)$$

where the state space vector $X^T = [S_o X_1 S_1 X_2 S_2]$ is comprised of the concentrations of: soluble organics S_o, mg/l; acidogenic bacteria X_1, mg/l; substrate for acidogenic bacteria S_1, mg/l; methanogenic bacteria X_2, mg/l; substrate for methanogenic bacteria S_2, mg/l. The only measurable and controlled output is the specific biogas production rate *Q*, l/l.d, $y = [Q]$, *Q*(0)=0.04. The inputs are the dilution rate as control variable *D*, d^{-1} ($D \in [0,0.3]$) and the influent organic concentration as disturbance S_{oi}, g/l, ($S_{oi} \in [30,70]$). The vector of the parameters

$q^T = [\beta\, Y_p \mu_{1\max}\; k_{s1}\; k_1\; Y_1\; \mu_{2\max}\; k_{s2}\; k_2\; Y_2\; Y_b\; Y_g]$

consists of the coefficients β, d^{-1}, Y_p, mg/l, Y_b, mg/g and Y_g, l/mg, the maximal specific growth rate of acidogenic $\mu_{1\max}$, d^{-1} and methanogenic $\mu_{2\max}$, d^{-1} bacteria respectively, the yield coefficients Y_1, mg/mg and Y_2, mg/mg, the saturation k_{s1}, k_{s2}, mg/l and the decay k_1, k_2, d^{-1} coefficients for the corresponding bacteria. The initial state vector is $X(0)^{oT} = [10\ 0.36\ 0.18\ 15.66\ 0.18\ ]$. The nominal parameter values are:

$q^{oT} = [1\ 2\ 0.4\ 1\ 0.02\ 0.006\ 0.4\ 1\ 0.02\ 1.1\ 40\ 1]$.

The problem is to develop neuro-fuzzy process models using MATLAB that account for the plant uncertainties due to variations in the inputs *D* and S_{oi}, in the initial states $X^T(0)$ and in the nominal plant parameters q^{oT} as result of ambient influences, microorganisms state, etc. The data for modeling is collected via simulation experiments after determination of the significant sources (initial states and parameters) of dominating influences via a sensitivity analysis.

The plant nonlinearity is studied by simulation of model (1). The step responses of the biogas production rate *Q* to equal incremental step changes of Δ0.05 of the input *D* within the range 0-0.3 when $S_{oi} = 50$, g/l is are shown in Fig.1. The influence of the different initial substrate inputs $S_{oi} = 30,40,50,60$ is given in Fig.2.

The problem is to develop neuro-fuzzy process models using MATLAB accounting for the plant uncertainties due to the change of the operation point along the nonlinear characteristic as a result of variations in the inputs *D* and S_{oi}, in the initial states $X^T(0)$ and in the nominal plant parameters q^{oT}, that reflect the ambient

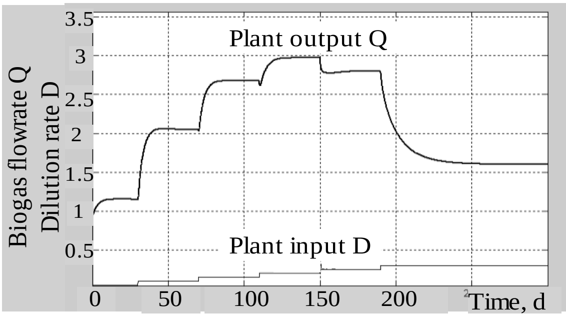

Fig.1. Sequence of plant step responses

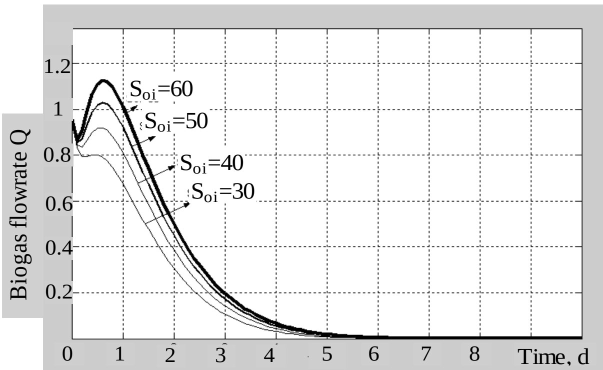

Fig.2. Plant output *Q* for different input S_{oi}

influences, the state of the microorganisms, etc.

The solution of the problem requires the accomplishment of the following tasks:

- determination of the significant sources (initial states and parameters) of dominating influences via sensitivity analysis;
- planning and carrying out of simulation experiments on developed simulation model based of (1) accounting for the variations in D, S_{oi} and significant $X^{T}(0)$ and q^{oT};
- development of a neural model by design and training of an ANN;
- development of a neuro-fuzzy model by design and training an ANN, representing Sugeno fuzzy model;
- simulation investigations for model accuracy assessment and comparison between the two models.

III. Sensitivity Analysis and Simulation Experimentation

The sensitivity analysis allows excluding correlated and insignificant parameters thus simplifying the plant model. Here it is based on the dimensionless sensitivity functions [3], [4]:

$$T_{Qq_k}(t) = \left.\frac{\partial \ln Q(t)}{\partial \ln q_k}\right|_{q_k^o}, \quad T_{QX_j(0)}(t) = \left.\frac{\partial \ln Q(t)}{\partial \ln X_j(0)}\right|_{X_j^o(0)} \qquad (2)$$

The significant simulated sensitivity functions are shown in Fig.3.

The sensitivity functions are local properties [4], so the conclusions deduced below by graphical sensitivity analysis depend on the nominal plant parameters and the initial conditions accepted:

1. The significant parameter and initial conditions set is determined for sensitivity functions greater than 1 (in the range of Q).
2. Parameters and initial conditions with similar influences on Q can be equivalently represented by one of them.

Thus the most significant parameters are Y_p, $X_1(0)$, $S_0(0)$ and β. They all are proportional and have a similar influence on Q. Since Y_b has a similar influence on Q like Y_p but influences less number of state variables and considering that only Q can be measured, so, variations in Y_b can reflect all possible influences in simulation experimentation for providing of realistic input-output data for neuro-fuzzy modelling. All the rest of the parameters and the initial conditions are considered to retain nominal values. Their variations have negligible impact on the biogas production rate or the impact of their variations can be reflected by another parameter changes. Simulation experiments for collecting of data for Q are carried out on a developed Simulink model based on (1), considering D=0.05;0.1;0.15,

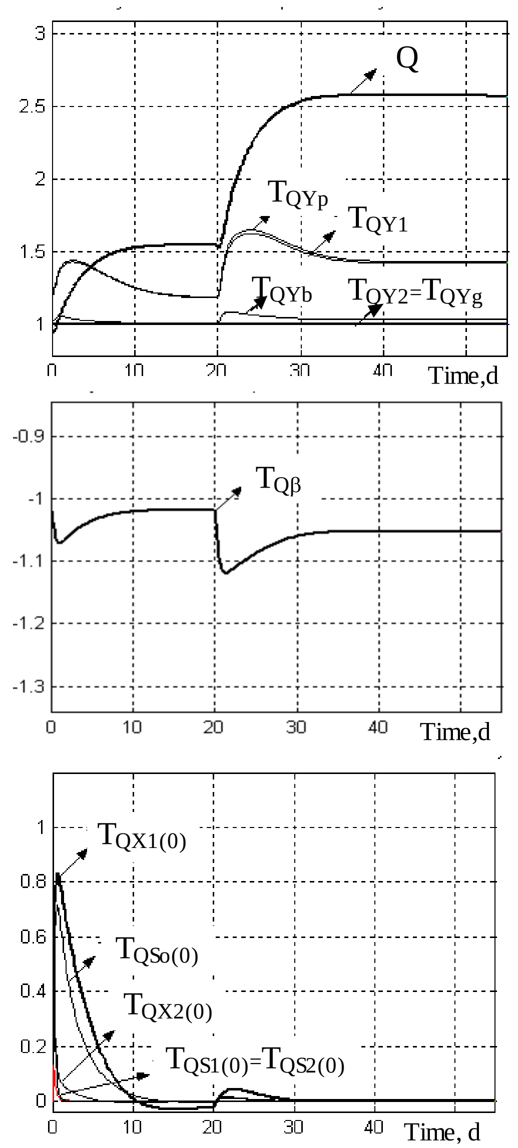

Fig.3. Sensitivity functions with respect to yield coefficients, kinetic coefficient β and initial conditions

S_{oi}=40;50;60 for each value of D and Y_b=30;40;50 for fixed D and S_{oi}.

IV. Design of Neural and Neuro-fuzzy Models

The first model developed is a two-layer feedforward neural network (NN) model with 7 hidden neurons and logsig activation functions for prediction in one time-step ahead, designed and trained using Neural Networks Toolbox of MATLAB and backpropagation method. The training data consists of 2 inputs [DN_i; QN_i] - 1 output QN_{i+1}, obtained from Q and D from the simulation experiments, shown in Fig.4, after normalization in the range [0, 1].

The final weighting matrices W_l and the bias vectors B_l of the two layers, l=1,2, obtained are:

$$W1 = \begin{bmatrix} -37.44 & -73.31 \\ -10.22 & -35.25 \\ 1.23 & 6.92 \\ 10.52 & 35.93 \\ 414.36 & -345.28 \\ -10.75 & -37.31 \\ -3.64 & -9.98 \end{bmatrix},$$

$$W2 = [-6.18\ 466.94\ 6.93\ 14.96\ -0.3\ -448.96\ 2.62]$$
$$B1 = [94.12\ 33.17\ -1.58\ -31.38\ -13.8\ 35.08\ 3.44]$$
$$B2 = [-18.09]\ .$$

In order to simplify the NN model an ANN, representing a Sugeno fuzzy logic (FL) model, is designed and trained using the same normalized training data and ANFIS (Adaptive Neuro-Fuzzy Inference System) from Fuzzy Logic Toolbox of MATLAB. For "and" and "or" operations are used the "min" and the "max" operators, the defuzzification method is the weighted average –wtaver. After a cluster analysis 3x2 Gaussian MFs for DN_i and QN_i respectively and 6 rules are initialized. The output functions are linear of the type $QN_{i+1}=a.DN_i+b.QN_i+c$. By training the Sugeno ANN the parameters of the input MFs and the coefficients in the linear outputs of each rule are tuned to reach the final goal –mean squared error below 0.001. The final MFs are shown in Fig.4. The model surface is highly nonlinear as seen from Fig.5. The tuned rule base is:

1. **If** DN_i is MF1 **and** QN_i is MF1
 then QN_{i+1}= 5.01DN_i + 0.48QN_i + 2.04
2. **If** DN_i is MF1 **and** QN_i is MF2
 then QN_{i+1}= -4.65DN_i - 0.92QN_i + 3.13
3. **If** DN_i is MF2 **and** QN_i is MF1
 then QN_{i+1}= 13.33DN_i + 1.76QN_i - 8.74
4. **If** DN_i is MF2 **and** QN_i is MF2
 then QN_{i+1}= 0.13DN_i + 1.76QN_i - 0.47
5. **If** DN_i is MF3 **and** QN_i is MF1
 then QN_{i+1}= 2.66DN_i + 1.27QN_i - 3.03
6. **If** DN_i is MF3 **and** QN_i is MF2
 then QN_{i+1}= 2.27DN_i + 1.82QN_i - 3.42

The simulated responses of the NN model Q_{nn} and of the neuro-fuzzy model Q_f as well as the relative modelling errors E_{nn} and E_f are shown in Fig.5. The Sugeno FL NN model is both more simple and accurate as seen from Fig.6.

V. CONCLUSIONS

The main contributions are the development and comparative study of two "black-box" models of the biogas production rate in anaerobic digestion of organic waste in waters, based respectively on classical feedforward backpropagation and Sugeno FL NNs using MATLAB. The models tackle plant uncertainty related to variable organic loading, ambient influence and microorganisms state. The models can find application in process prediction, optimization and control.

REFERENCES

[1] DT. Hill, and CL. Barth, "A dynamical model for simulation of animal digestion," J. Water Pollution Contr. Fed 10, 1977, pp. 2129-2143.

[2] D. Dochain, and P. Vanrolleghem, Dynamical Modelling and Estimation in Wastewater Treatment Processes. IWA Publishing, UK, 2001, p. 312

[3] T. G. Müller, N. Noykova, M. Gyllenberg, and J. Timmer, "Parameter identification in microbial population dynamics models of anaerobic wastewater treatment," J. Math. Biosci. 177-178 2002, pp.147-160.

[4] N. Noykova, T. G. Müller, M. Gyllenberg, and J. Timmer, "Quantitative analysis of anaerobic wastewater treatment processes: identifiability and parameter estimation," J. Biotech. Bioeng., 78 (1), 2002, pp. 89-103.

[5] I. Simeonov, V. Momchev, and D. Grancharov, "Dynamic modeling of mesophilic anaerobic digestion of animal waste," J. Wat. Res. 30 (5), 1996, pp. 1087-1094.

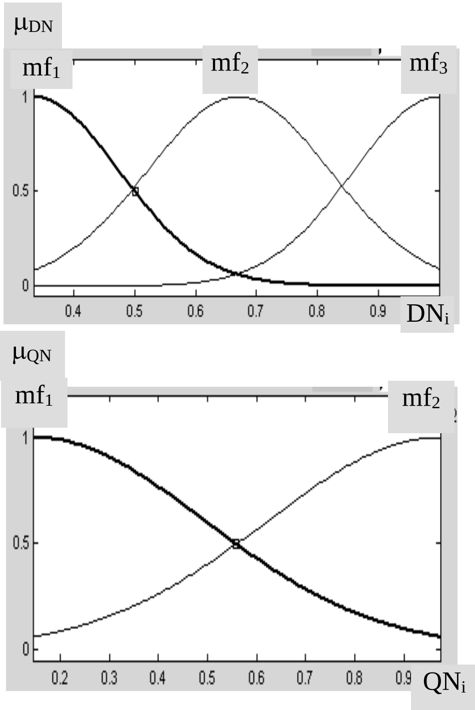

Fig.4. Membership functions for the input variables *DN* and *QN*

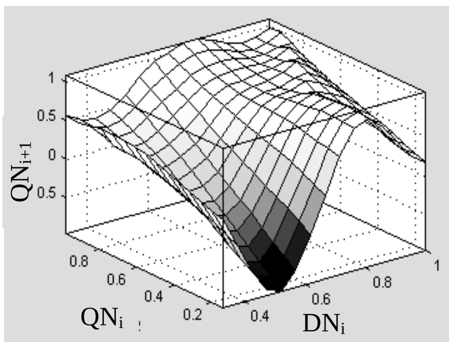

Fig.5. Input-output model surface

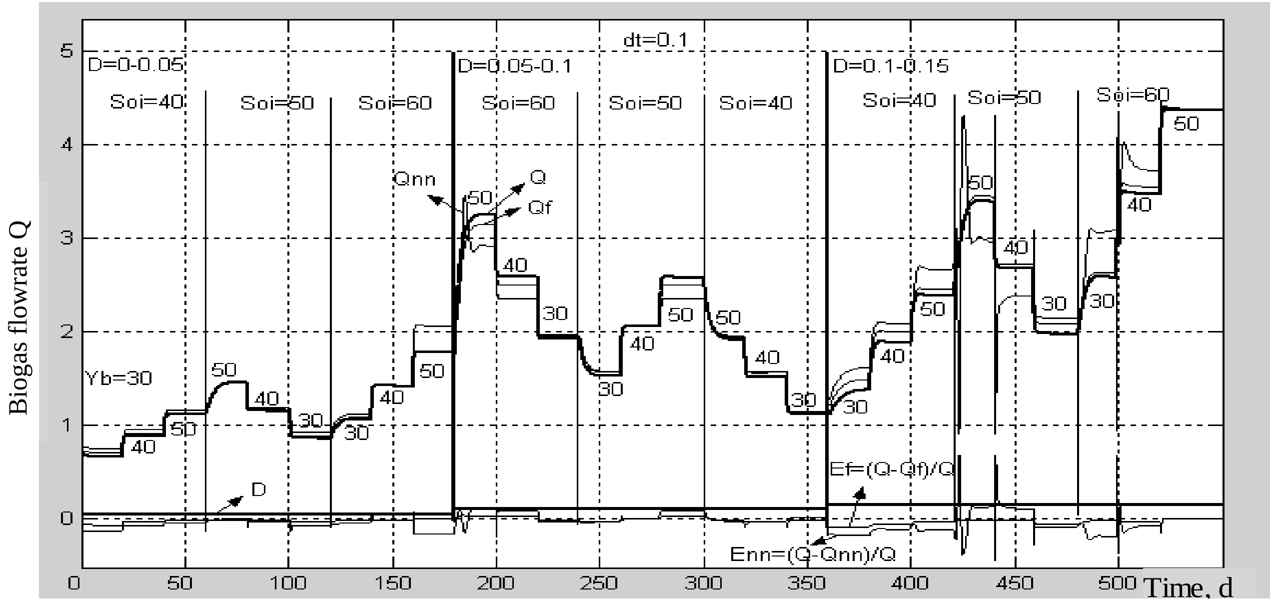

Fig.6. Simulink simulation experiments using plant model (1), NN and FL plant models

[6] I. Simeonov, “Modeling and control of anaerobic digestion of organic waste,” J. Chem. Biochem. Eng. Q. 8 (2), 1994, pp. 42-52.

[7] O. Bernard, Z. Hadj-Sadok, and D. Dochain, “Advanced monitoring and control of anaerobic wastewater treatment plants: II - dynamic model development and identification,” Proc. of the 5th Int. Symp. WATERMATEX, Gent, Belgium, September 2000. pp. 3.57-3.64.

[8] S. Yordanova, and N. Noikova, “Influence of perturbations on the waste water treatment process,” J. Chem. Biochem. Eng. Q. 10 (1), 1996, pp. 9-14.

[9] T. J. Ross, Fuzzy Logic With Engineering Applications. McGraw Hill, Inc., 1995.

[10] J.-S.R. Jang, C.-T. Sun, and E. Mizutani, Neuro-Fuzzy and Soft Computing. A Computational Approach to Learning and Machine Intelligence. MATLAB Curriculum Series, Prentice-Hall, Inc., N.J., 1997. ISBN 0-13-261066-3.

[11] S. Yordanova, “Fuzzy two-level control for anaerobic wastewater treatment,” Proc. of the 2-nd Int. IEEE Conf. “Intelligent Systems”, Varna, June 22-24, 2004, Sofia, Bulgaria 2004, pp. 348-352.

[12]`S. Yordanova, R. Petrova, and V. Mladenov, “Neuro-fuzzy control for anaerobic wastewater treatment,” WSEAS Trans. on Systems, 3, 2004, pp. 724-129, ISSN 1109-2777.