Neural Network Control in a Wastewater Treatment Plant

Miguel A. Jaramillo[1]; Juan C. Peguero[2], Enrique Martínez de Salazar[1], Montserrat García del Valle[1]

([1])Escuela de Ingenierías Industriales. ([2])Centro Universitario de Mérida. Universidad de Extremadura. Avda. de Elvas s/n. 06071 Badajoz SPAIN.
e-mail: miguel@unex.es

Abstract. Control in a Wastewater Treatment Plant is a task mainly performed by technicians because the complexity of the dynamics of the biochemical processes performing the wastewater treatment and the delays in the measurement of several variables describing the system dynamics prevent from using classical automatic regulators. A Neural Network that learns the control actions carried out by the technicians controlling the plant is developed and then used as an automatic controller in a simulation of the plant dynamics.

1 Introduction

Wastewater Treatment Plant are facilities where municipal and industrial wastewaters are processed to eliminate as much pollution as possible. When these waters enter the plant large particles are rejected and greases and oils are skimmed off in a first stage. In a second one organic pollution is removed with a biological treatment where bacteria "eats" the pollutants in the water reducing its concentration to an appropriate level. Then this water may be released to a natural stream. This second stage may be divided into two different processes. The first one is performed in an aeration tank where the incoming polluted water is mixed with a sludge made up of bacteria. After they have "eaten" most of the organic matter water and sludge are driven to a settler tank where the second process is performed: the separation of these two components. The sludge flows downwards while the water stays at the top of the tank and flows over an overflow weir to be released to natural streams. The sludge is withdrawn from the tank and then divided into two streams: one is driven to the aeration tank to keep the sludge concentration at an appropriate level while the other is eliminated.

From the preceding description it may be inferred that the sludge feedback must play an important roll in the biological process control. In fact it may be considered as the most important control parameter of the plant dynamics. It allows the sludge concentration in the aeration tank to be kept at an appropriate level to ensure an adequate plant operation.

As the plant dynamics must be described by biological processes, the formulation of a mathematical model is very difficult. Therefore a simplification must be applied in order to obtain a computational treatable structure along with a reasonable understanding of the plant dynamics. Fortunately a good approximation to the behavior of the plant may be obtained with the description of only the aeration tank, the main element of the whole system. The equations describing this reduced model are nonlinear and the use of classical linear control techniques is not possible. Moreover some of the system variables need several days to be obtained and, therefore, automatic controllers may not be included, as they need online variables values (or at least with small time delays) to operate. So the plant must be controlled by technicians that adjust the plant parameters. As they have only delayed values of the plant variables they must rely on their own experience to adjust the values of those parameters. This experience may be complemented with the direct observation of some properties of the fluids in the tanks, as their apparent dirtiness, that allows the operator to suppose an approximate concentration of bacteria and pollutants. The operator experience is then based on the study of the effects his actions produce once the actual values of the plant variables (several days delayed) have been obtained.

As wastewater treatment plant technicians provide a very effective control, the learning of their experience may define a very effective tool to provide an automatic control of the plant. Nevertheless, as these plants are continuously working, it is not possible to test the performance of such a controller, so it will be necessary to simulate both the plant and the control. This is the aim of the present work: the definition of the model of a wastewater treatment plant controlled by a system that reproduce the operator's knowledge. As this knowledge is based on personal experience it must be learned with artificial intelligence techniques. We have selected Neural Networks as that tool because the operator's experience is represented by a time series and they have proved their capabilities to learn a dynamical system behavior from its time evolution [1]-[2]-[3]-[4]. As all the information concerning the plant dynamics is available at the training time, the proposed model may learn the operator's actions relating then to the corresponding system variables. In this way the operator's knowledge is associated with the plant variables. Once the neural network has been trained it will be included in the plant model to test its capabilities in simulation.

2 Plant description

A diagram of the model to be used in this work [5]-[6] appears in Figure 1.-. It is focused on the aeration tank, where the main processes of the plant take place. Parameters defining a real plant also appear. In this model the input variables are the "Influent Flowrate" *Qf* and the "Influent Pollution Concentration" *Sf*. The output variables are the "Sludge Concentration" in the aeration tank *Xva* and the "Output Pollution Concentration" *Se*. The "Sludge Concentration" in the settler, *Xvu*, is assumed to be a process parameter that will be provided at every time step. This parameter defines de concentration of the sludge in the settler, a portion of which will be recirculated to the aeration tank to keep the sludge concentration at an appropriate level. This recirculation is defined by a flowrate, *Qr*, that is adjusted by the operator supervising the plan dynamics. It is represented by a parameter, *r*, that provides the ratio between this "Recirculation Flowrate" and the "Influent Flowrate".

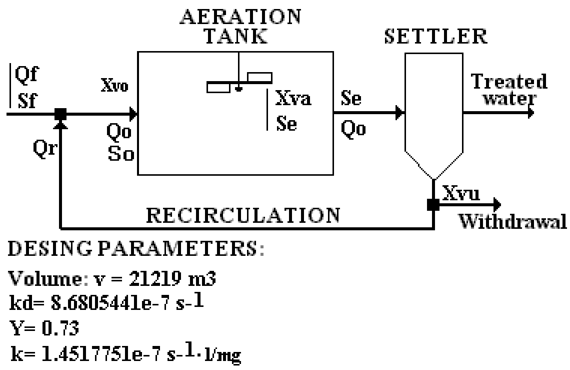

Fig. 1. Simplified diagram of the wastewater treatment plant.

The system dynamics may be defined by a set of differential equations of the variables *Se* and *Xva*. No dependence on the temperature or the aeration will be considered, their values will be assumed to be optimal, a supposition that is quite close to reality because the great volume of the tanks in the plant ensure a quite constant value for temperature, while the aeration may be set to an adequate level by air pumping. So the proposed model is described by the following equations [5]-[6]:

$$\frac{dSe}{dt} = -\frac{Qf}{V} \cdot Se - k \cdot Xva \cdot Se + Sf \cdot \frac{Qf}{V} \tag{1}$$

$$\frac{dXva}{dt} = -\left[\frac{Qf}{V} \cdot (1+r) + kd\right] \cdot Xva - \frac{Qf}{V} \cdot Y \cdot Se + \frac{Qf}{V} \cdot r \cdot Xvu + \frac{Qf}{V} \cdot Y \cdot Sf \tag{2}$$

Recirculation is defined by:

$$r = \frac{Qr}{Qf} \cong \frac{Xva}{Xvu - Xva} \tag{3}$$

The equality will be valid only at equilibrium, where the plant is assumed to work.

3 Control of the recirculation parameter

As it has been stated previously the plant control is performed adjusting the value of the recirculation parameter *r*. To provide an easier manipulation it has been expressed as the sum of two elements: $r=ro+\delta r$. The first one is fixed and define a mean value of the recirculation, while δr represents the oscillations around that value. They have been selected so that δr has only values between +1 and –1 to ensure an optimal neural behavior, as it will be seen later. *ro* has been set to ro=1.74, a value obtained from the study of the time series defining *r* in the real plant used in this work.

As this parameter is adjusted by the plant operator, whose experience is to be learned by a neural network, the time series describing its evolution may be used to perform the network training along with all the other time series of the plant variables and parameters. So, the neural network may learn the value the operator gave to the recirculation relating it to the values of the plant variables and parameters obtained at the same time. Al these data have been taken from the activity reports of a real Wastewater Treatment Plant [7]. In this way the operator experience may be related to the plant variables and parameters providing a model of his behavior that may be used as a controller in simulation, where direct values of the variables may be used to control the system dynamics.

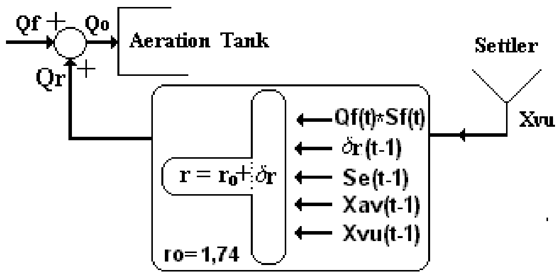

Fig. 2. Control of the recirculation constant

So the variables and parameters used as network inputs are *Sf(t)Qf(t)*, *Xva(t-1)*, *Se(t-1), Xvu(t-1)* and δ*r(t-1),* while the output is δ*r(t)* (Figure 2.-). We have considered the product of the input flowrate and its concentration instead of their separated values because that product represents the value of the total pollutant mass, which have a more direct influence on the sludge evolution than the separated values. Several simulations carried out both with the product and with their separated variables have proved that better results were obtained with the fist configuration.

4 Neural Network structure

The selected Neural Network is a Multilayer Preceptron [4] with two hidden layers. The input one has 5 inputs (those defined in the previous section as inputs to the network). The first hidden layer has 15 elements while the second one has 10. The output layer has only one neuron whose output is the recirculation correction δ*r* , the network response.

Every neuron gives an output that is function of the weighted sum of all the outputs of the preceding layer and a bias term. This function is a saturating one usually known as sigmoid because of its s-shape. So the output of the j-th neuron in the i-th layer is:

$$y_j^i = \sigma\left(\sum w_{jk} y_k^{i-1} + I_j\right) \tag{4}$$

Feedback has not been considered in the neuron model because, as feedforward Neural Networks are universal approximators [8], it is not necessary to include it to identify the time evolution of a dynamic systems. Moreover the inclusion of feedback between neurons will introduce a destabilization factor that must be studied so that a demonstration of the network stability must be provided before using it in system identification.

The output functions for neurons in the two hidden layers are sigmoids that provide a value between 0 and 1:

$$\sigma(x_j) = \frac{1}{1+e^{-x_j}} \tag{5}$$

while the neuron in the output layer has an arctangent to provide the desired variation of δr between +1 and –1.

The bias factor I_i has been included in the neuron model in order to obtain a good fitting between the weighted sum of the neuron inputs and the neuron output function resolution so that the value of the sum will lay in that zone of this function where a higher resolution is provided. In this way both weights and bias will be adapted during the learning process [9].

In many applications of neural networks the bias term is not included and the necessary fitting between the weighted sum and the output function must be provided by an appropriate normalization of the input data. Usually this is a trial and error task where several procedures must be tested to obtain that providing the best result. This process needs a previous analysis of the input data to decide the normalization that may provide a better fitting between the input data and the output function. In this way a good result may always be obtained although it may not be assured that it is the best possible. Moreover, when data sets different form that used to perform the normalization process are presented to the net, a worse result may be obtained if the new data have a distribution different enough from those used to define the normalization, a fact that will make difficult the selection of a satisfactory process that could be used with a wide range of input data. The inclusion of the bias term may provide a compensation for those deviations if it is given an appropriate value. This value must be obtained along with the input weights during the training process of the network. In this way the presence of the bias term in the neuron model will provide it with the capability to adapt the output function to a wide range of values of the weighted sum the neuron performs.

Although the bias term has been included to complement the normalization process of the input data, it may be considered in all the layers to obtain a better adjustment between the neuron output and its inputs. So, as I_i may be assumed as a function parameter, (4) and (5) may be rewritten as:

$$y_j^i = \sigma\left(\sum w_{jk} y_k^{i-1}\right) \tag{6}$$

$$\sigma(x_j) = \frac{1}{1+e^{-(x_j + I_j)}} \tag{7}$$

Obviously the inclusion of the bias term does not mean that the normalization of the input data must not be done, since they usually have values that are not fitted to those the network can receive. But now the normalization process must only provide an adaptation between the input data set and the network input. The subsequent adaptation of the bias term during the learning process will provide the necessary adjustment between then to obtain the best performance.

Following this procedure we have normalized all the input data to fit them to the network input. We have divided each element of the time series defining the time evolution of each input by its mean value, a value rough enough to be used as normalization constant by any set of data that may be used later, including those obtained by simulation.

5 Training of the Neural Network structure

The network training was performed with a “Backpropagation” that used the Levenberg-Marquardt algorithm [4]. In order to have an easier implementation of the learning algorithms we considered the neuron model described by (4) and (5) so that the bias term may be assumed as an weight that always receives an input that is equal to 1. In this way algorithms provided by commercial software may be used easily. We have performed our simulations with MATLAB.

Two years of activities in the real plant [7] were considered: 1998 for training and 1999 for simulation. These data were shared out into two groups, one for the first three and last three months of the year (which we named wet months) and the other for the remaining six (which we named dry). This distribution try to ensure a more or less uniform temperature distribution through the six months that form each group to fulfill the aforementioned supposition that the temperature is kept to a constant value. So two different networks were generated. The plant model used equations (1) to (3) with the parameters of the same plant that provided the time series. Their values appear in Figure 1.-.

In the training process all the data have been taken from the corresponding time series provided by the activity reports of the real plant. In the simulation some of them were taken form that reports (*Sf(t)Qf(t)*, as actual input data, and *Xvu(t-1)* as a plant parameter) while others have been provided by the simulation (*Xva(t-1), Se(t-1)* and $\delta r(t-1)$)).

6 Results and comments

Once the neural network has been trained it was added to the plant model defined by equations (1)-(3) to provide a value for δr and the simulation of the plant with the neural controller was performed. To check the good behavior of the proposed controller, the simulation was repeated with δr obtained from the time series of the real plant. In Figures 3.- and 4.- the results obtained for two wet months, April and December, are presented. Very similar results were obtained with dry months.

First of all we can see that the neural controller provides a very effective regulation of *Xva* and *Se.* They are kept at very good levels, very close to those the plant was designed for: about 2000 mg/l for *Xva* (fluctuations between 1500 a 2500 mg/l are also allowed) and about 7.5 mg/l or less for *Se*. It can be seen that the output pollution concentration is very similar to that obtained when the control is performed by the time series describing the operator's actions, what proves the effectiveness of the neural network at learning the operator's skills. Differences are bigger for the sludge concentration but both dynamics are very similar. Moreover the simulation with the neural controller provides values that are closer to those assumed in the plant design than those provided by the time series of the recirculation. This is not a surprising fact because while the operator must adjust that parameter with delayed measures of the plant variables and subjective observations, the neurocontroller uses precise values. So the combination of the operator's experience with direct measurements of the plant variables provides a more effective regulation of that variables. If it were possible to measure that variables with a little delay in real plants the defined controller might be used to provide an automatic control, because as the time evolution of the plant variables is quite slow (it takes several hours to notice the effect of a modification in the input values) time delays smaller than an hour will allow an almost real time control. Unfortunately devices providing those fast measures are not used in real plants because they are too expensive.

References

[1] Narendra, S. N. and Parthasarathy, K. "Identification and Control of Dynamical Systems Using Neural Networks", IEEE Trans. on Neural Networks, Vol. 1, No.1, March 1990, pp 4-27.

[2] Simpson P. K. (Editor) "Neural Networks Applications", IEEE Technology Update Series, 1996.

[3] Hrycej T., "Neurocontrol, Towards an Industrial Control Methodology", Wiley Interscience, 1997.

[4] Norgaard, M., Ravn, O., Poulsen, N. K. And Hansen, L. K. "Neural Networks for Modelling and Control of Dynamic Systems", Spinger-Verlag 2000.

[5] Ramalho R. S. "Tratamiento de Aguas Residuales". Reverte, 1996.

[6] Olsson, G.; Newell, B. "Wastewater Treatment System". IWA, 1999.

[7] E.D.A.R. Aguas de Mérida S. L., "Operation and Maintenance Reports". 1998 and 1999.

[8] Hornik, K., Stinchcombe M. and White, H. "Multilayer Feedforward Networks are Universal Approximators". Neural Networks, Vol. 2 (1989), pp. 359-366.

[9] Chen, C.-T., Chang, W.-D., "A Feedforward Neural Network with Function Shape Autotunig". Neural Network Vol. 9 (1996), No. 4, pp 627-641.

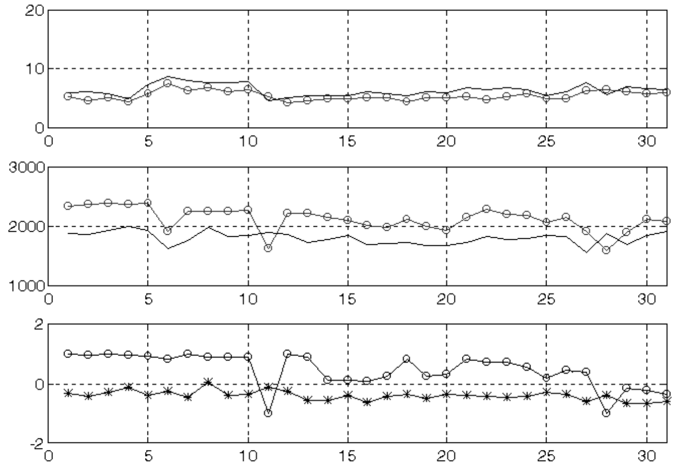

Fig. 3. April 1999. Top to down *Se* (mg/l), *Xva* (mg/l) and δ*r* (no dimension). Line with “o”: data obtained with the neurocrontroller. Line without “o”: data obtained from operator.

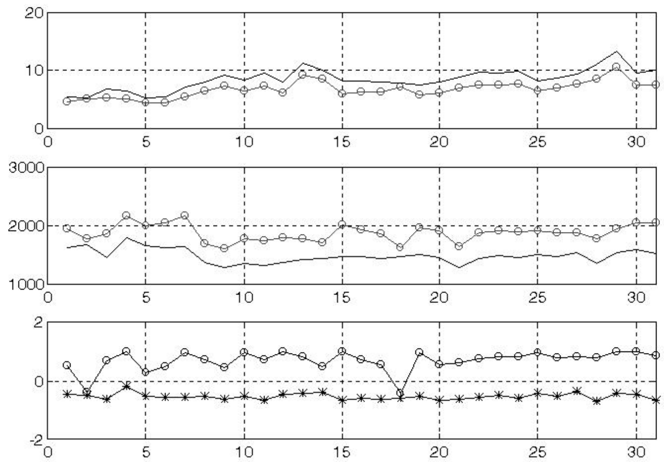

Fig. 4. December 1999. Top to down *Se* (mg/l), *Xva* (mg/l) and δ*r* (no dimension). Line with “o”: data obtained with the neurocrontroller. Line without “o”: data obtained from operator.