DRINKING WATER TREATMENT: A NEURAL NETWORK MODEL FOR COAGULATION DOSING

B. Lamrini*, A. Benhammou*, A. Karama* and M-V. Le Lann**

** Laboratoire d'Automatique et d'Etude des Procédés, Faculty of Sciences Semlalia, PoBox:2390, 40000-Marrakech, Morocco*
E-mail :{lamrin.b, benhammou, karama@ucam.ac.ma}
*** LAAS/CNRS, 7, Avenue du Colonel Roche, 31077-Toulouse cedex 4, France and INSA, DGEI, 135,Avenue de Rangueil, 31077-Toulouse cedex 4, France*
E-mail: mvlelann@laas.fr
All correspondence should be addressed to: benhammou@ucam.ac.ma

Abstract: The aim of this paper is to present the development and validation of a neural network model for on-line prediction of coagulant dosage from raw water characteristics. The main parameters influencing the coagulant dosage are firstly determined via a PCA. A brief description of the methodology used for the synthesis of neural models is given and experimental results are included. The training of the neural network is performed using the Weight Decay regularization in combination with Levenberg-Marquardt method. The simulation results of neural model compared to a linear regression model are illustrated with real data.

Keywords: Coagulation process; drinking water treatment; neural networks; weight decay regularization.

1. INTRODUCTION

The use of artificial neural networks for process modelling and control in the drinking water treatment is currently on the rise and is considered to be a key area of research. The coagulation process which requires the addition of chemical coagulant is the critical process in the drinking water treatment. The control of a good coagulation is essential for maintenance of satisfactory treated water quality and economic plant operation. Basically, coagulant dosage is chosen empirically by operators based on their past experience, laboratory jar-testing and various information on water quality parameters. The jar-test apparatus simulates mixing, flocculation, setting, and a single test may take about one hour to be performed. Disadvantages associated with jar-testing are that regular samples have to be taken requiring manual intervention and operators can make manually in raw water quality. There is no mechanistic model describing the coagulant dosage related to the different variables affecting the process. Consequently, there is a need for a fast and reliable method for determining the required coagulant rate which can be used instead of the jar-test analysis.

The purpose of this paper is to highlight the utility of artificial neural networks in drinking water treatment in particular coagulation modelling and control. Process data can be used directly to represent input-output process relationships. Neural networks proved to be extremely flexible in representing complex non-linear relationships between many different process variables (Cybenko, 1989). They do not require any a priori precise knowledge on the relationships of the process variables. Various applications of these models have been recently reported in the drinking water treatment industry (as examples: the forecasting of drinking water (Canu *et al*., (1990, 1997)), the prediction of the coagulant dosing (Gagnon *et al*., 1997; Agdar *et al*., 1996; Evans *et al*., 1997)).

This investigation aims to develop a neural model for the on-line prediction of optimal coagulant dosage from raw water characteristics. Previous researchs show the efficiency of a such approach using neural networks (Gagnon *et al*., 1997; Agdar *et al*., 1996).

This paper is organised as follows. First, the drinking water treatment process is described . Then a brief description of neural network is presented. In the subsequent section, the neural model for the on line prediction of the coagulant dosage is developed. Finally, experimental results are presented.

2. WATER TREATMENT PLANT DESCRIPTION

The plant of drinking water treatment concerned by this study is the drinking water treatment plant Rocade located in Marrakech. It provides water to more than 1,5 millions inhabitants. Raw water is extracted from the channel Rocade. In case of resource failure (raw, pollution...), the treatment plant takes the raw water from a pumping plant Takerkoust. 60% of city needs are assured by the treatment plant, the complement is brought by the underground resources (well, drilling…). It has a nominal capacity to process 1400 l/sec of water. The treated water is stored in two tanks and transported toward the water supply network.

The drinking treatment plant involves physical and chemical processes. Figure 1 presents a schematic overview of the various operations necessary to treat the raw water at the Rocade water treatment plant of Marrakech. The treatment consists in essentially of preliminary disinfection,, coagulation-flocculation, settling, filtration and final disinfection.

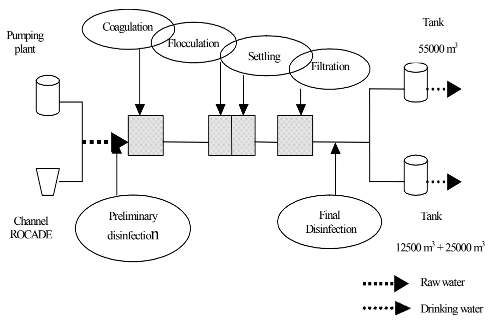

Fig. 1. Simplified synopsis of Rocade water treatment plant

3. OVERVIEW OF ARTIFICIAL NEURAL NETWORKS

Neural networks are known to be able to successfully represent complex functions in various fields. There are a wide variety of artificial neural networks existing in the literature, of which the feedforward structure is one of the most commonly used in modelling and control. Feedforward neural networks, such as the multilayer perceptron, consist usually of many simple processing elements arranged in layers as shown in Figure 2. Each element takes its input from the weighted sum of the outputs of the elements of the previous layer. This input is then passed through a nonlinear function, often called the activation function, to form the element's output (Hertz, 1991). In this study, a three layer feedforward neural network has been adopted: it has been shown that is a network with a single hidden layer (as shown in figure 2) can simulate any continuous function.

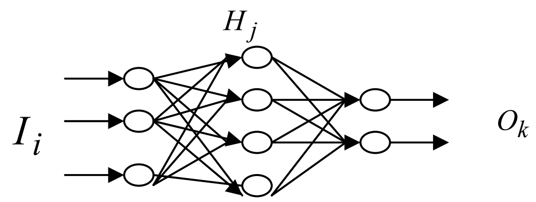

Fig. 2. Feedforward neural network

The neural network models consist of the following set of equations:

$$O_{k_k} = f_k\left(\sum_j w_{jk} H_j\right) \text{ and } H_{j_j} = f_j\left(\sum_i w_{ij} I_i\right) \qquad (1)$$

where O_k denotes the network outputs, H_j the hidden neurons outputs, I_i the network inputs, w_{ij} the weights between the input layer and the hidden layer and w_{jk} the weights between the hidden layer and the output layer.

Training of the neural network involves adjusting the weights w_{ij} and w_{jk} by using the backpropagation learning algorithm (Rumelhart & McClelland, 1986), so that the network emulates the non-linear function underlying the training data set. The network weights are adjusted by minimising the following criteria derived from the difference between real and neural outputs respectively t and O .

$$C = \frac{1}{N}\sum_{i=1}^{N}(t_i - O_i)^2 \qquad (2)$$

4. PREDICTION OF COAGULANT DOSAGE

The coagulation process involves many complex physical and chemical phenomena which are difficult to model using mechanistic and chemical phenomena traditional description. The coagulant dose ensuring optimal treatment efficiency has been shown experimentally to be non-linearly correlated to raw water characteristics which are usually available on line.

In the sequel, a neural software sensor for the prediction of coagulant dosage is developed in too stages. The Factorial Analysis in Principal Component method is firstly applied to determine the main parameters affecting the prediction of the optimal coagulant dosage. These parameters will be then considered as the input variables of the neural model for which the training algorithm will be performed.

4.1 Factorial Analysis in Principal Component

As far as the dimensional analysis is concerned, 9 describers of the raw water quality (temperature, pH, turbidity, Total carbonates, total suspend solid (TSS), oxydability, dissolved oxygen, conductivity and the coagulant dosage (passive parameter)) are used. A number of 89 samples have been used like individuals. Every sample underwent different physical and chemical analysis as well as to the jar-testing to determine the coagulant rate. In order to show the various relationships between variables and principal components, It is interesting to visualize this variables in the relationship circle. This representation allows to compare the behavior of a variable beside the other variable set (It is the case of turbidity (3) – TSS (4)). On the other hand, the following variables have an inverse behavior (diametrically opposite in the circle): turbidity and TSS with dissolved oxygen (5), oxydability (7) with dissolved oxygen.The total carbonates (8), situated to the center of the interrelationship circle, seems to have no effect on the system. it cannot be interpreted. The temperature (2) presents an independent behavior with to other variables. It wouldn't have to be eliminated in order to perfectly explain variations of the system. The same thing for conductivity and pH variable. The process of simplification would allowed to extract the following variables: temperature, pH, TSS, oxydability, dissolved oxygen and conductivity. Given that the survey consists to predict the coagulant dosage related to describers easily measurable on-line, we have keep solely the following variables: temperature, pH, TSS, dissolved oxygen and conductivity. The oxydability variable is eliminated because it is not yet measurable on-line.

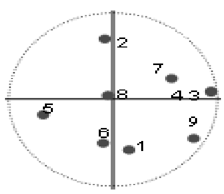

Fig. 3. Relationship circle.

4.2 Neural model training

The neural network used for modelling of the coagulant dosage is a MLP type. Note that the network inputs are the observed values of the retained raw water quality parameters.

For the determination of the architecture network, the pruning approach "*Weight Decay*" is used, starting from a relatively large network then removing connections in order to arrive at a suitable network architecture (Hinton, 1987; Cibas & Gallinari, 1999). This approach, allowing to eliminate the weak weights, consists in adjusting the weights using the new performance function instead of C defined in (2):

$$C'(w) = C(w) + \alpha\,\Omega(w)$$

where $$\Omega(w) = \frac{1}{n}\sum_{i=1}^{n} w_i^2 \qquad (3)$$

n is the number of weight network, α is a parameter that determines the importance of the two terms in the new performance function $C'(w)$. Using this performance functions will cause the network to have smaller weights and biases, and this will force the network response to be smoother and less likely to overfit. This method presents the advantage to be simple to implement, since the gradient of C' can be very easily calculated from the gradient of C and from network weights. It is sufficient to add the quantity αw to the gradient vector ∇C calculated by the Back-propagation algorithm:

$$\nabla C' = \nabla C + \alpha\, w \qquad (4)$$

5. RESULTS AND DISCUSSION

The experimental data of four years (2511 samples, from January 2000 to July 2003) have been used to establish the neural model as the basis of a coagulant dosing estimation software. For the conductivity and the oxygen dissolved, invalid data have been removed and missing ones have been replaced. Figure 4 shows the wide range of raw water quality describers that exist on the drinking treatment plant.

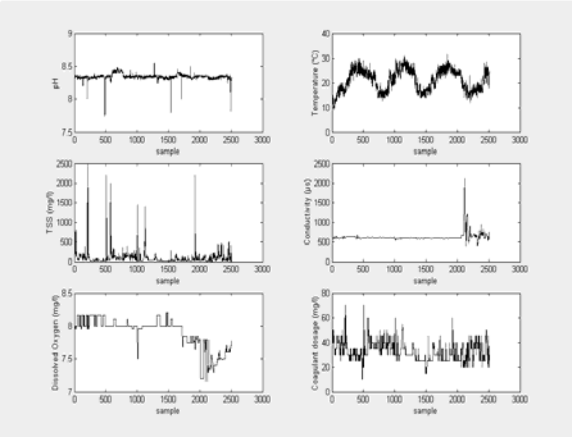

Fig. 4. Raw water characteristics.

75% of the global data is used for the network training and the remainder (25%) for validating the neural model. The obtained networks consist of five inputs, a single hidden layer with 18 sigmoidal neurones and one linear output. The neural model have been built using the regularisation method in combination with Levenberg-Marquardt training algorithm. It has 127 connections in the beginning of training. The advantage of this algorithm is that it provides a measure of how many weights of the network are efficiently used by the network. In our case, the final qualified network uses approximately 82 weights, out of the 127 total weights in the 5-18-1 network.

Figure 5 shows the validation of the neural model. We notice that the coagulant dosage computed with the neural network model is very smoothed to the real data. Consequently, the neural network generalizes well to new data. The obtained sum of the squared error is 0.007. Figure 6 shows the correlation between neural output and real data. The correlation coefficient computed on the validation set is equal to 0.94.

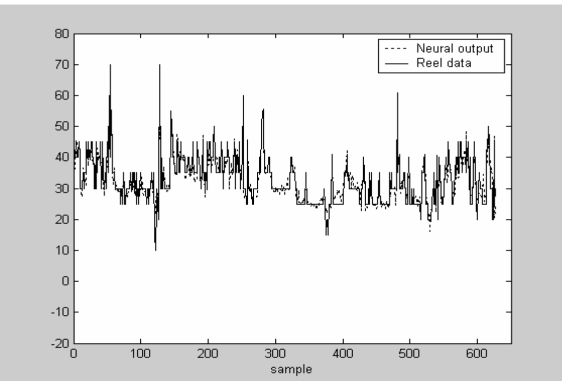

Fig. 5. Neural coagulant dosing rate with respect to the reel data on the validation data.

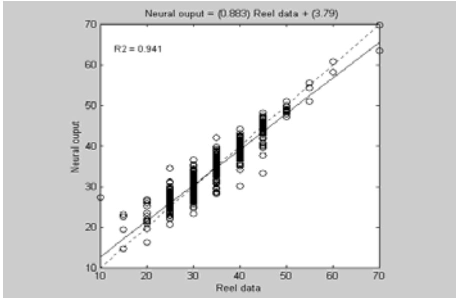

Fig. 6. Correlation between neural coagulant dosing rate and real data.

A linear regression model for the coagulant dosing rate has been also developed for the comparison with neural network results. The correlation coefficient between linear model and real data on the validation set is equal to 0.48. It is much smaller than the one obtained with to the neural network (0.94). Figure 7 shows the output of the linear model on the validation data. We clearly see that the prediction accuracy is inferior to the one of neural network model. Furthermore, as is shown in the same figure, we can deduce that the linear model also computed erroneous values (a negative rate)..

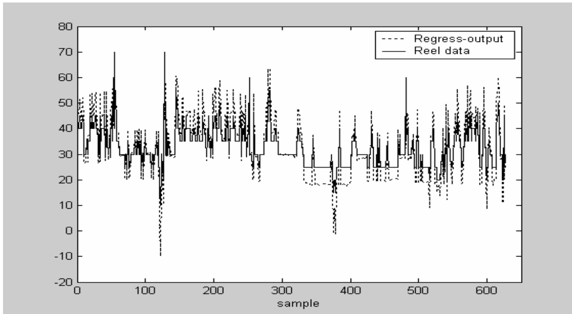

Fig. 7. Comparison of the regression and the real output on the validation data.

6. CONCLUSION

This paper has addressed the development of neural networks model for the prediction of the coagulation dosage for the Rocade water treatment plant. A large data bank, obtained over many years of operation, has been used to develop this model. In conclusion, we showed in this paper that the coagulant dosing is non-linearly correlated to the raw water characteristics such as TSS, temperature, pH, conductivity and dissolved oxygen. Experimental results using the data raw water plant showed the efficiency and soundness of this approach. The performance of the network depends on the quality and the completeness of data provided for training the system.

REFERENCES

Adgar, A., C.S. Cox., P.R. Daniel. and A. Lowdon (1996). Chemical dosing philosophies for a water treatment plant: results of some pilot plant experimentation. *In Proceeding of the IEE Conference Publication*, 427(2), 1052-1057.

Canu, S., X. Ding. and Y. Granvalet. (1997). *Statistiques et Méthodes Neuronales* : Une application des réseaux de neurones pour la prévision à un pas de temps. S. Thiria, Y. Lechevallier,O. Gascul et S. Canu (Ed), Chapitre 7, 120-131, Dunod, Paris.

Canu, S., R. Sobral and R. Lengellé. (1990). Formel neural network as an adaptative model for water demand. *Proceedings of INNC'90*, Kluwer Academic Publishers, 1, 131-136.

Cybenko, G. (1989). Approximation by superposition of a sigmoid function. *Math. of Control Signals and Systems*, 4(2), 303-314.

Evans, J., h.C. Enoc., M. Johnson. and P. Williams. (1998). Intelligent based auto-coagulation control applied to a water treatment works. *In Proceedings of International Conference on Control*, 141-145.

Gagnon, C., B.P.A. Grandjean. and J. Thibault. (1997). Modelling of coagulant dosage in a water treatment plant. *Artificial Intelligence in Engineering*, 11, 401-404.

Gallinari, P. and T. Cibas. (1999). Practical complexity control in multilayer perceptrons. *Signal Processing*, 74, 29-46.

Hertz, J., A. Krogh. and R.G. Palmer. (1991). Introduction to the Theory of Neural Computation. Addison-Wesley.

Hinton, G.E (1987). Learning Translation Invariant Recognition in Massively Parallel Networks. *Proceedings PARLE Conference on Parallel Architectures and Languages Europe*, J. W. de Bakker, A. J. Nijman, P.C. Treleaven (Eds.), 1-13, Springer-Verlag, Berlin,.

Mackay, D.J.C. (1992a). Bayesian interpolation, *Neural Computation*, 4, 415-447.

Mackay, D.J.C. (1992b). A pratical Bayesian Framework for backpropagation Networks, *Neural Computation*, 4, 448-472.

Rumelhart, D.E.and J.L. McClelland. (1986). Parallel distribution processing: exploration in the microstructure of cognition, 1, MA: MIT Press, Cambridge.