ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Environmental Modelling & Software xx (2006) 1–6
Environmental Modelling & Software
www.elsevier.com/locate/envsoft

Short communication

Simulation of a coke wastewater nitrification process using a feed-forward neuronal net

I. Machón [a], H. López [a], J. Rodríguez-Iglesias [b], E. Marañón [b,*], I. Vázquez [b]

[a] Department of Electrical Engineering, Computers and Systems, University of Oviedo, 33203 Gijón, Spain
[b] Department of Chemical and Environmental Engineering, Higher Polytechnic School of Engineering, University of Oviedo, 33203 Gijón, Spain

Received 3 May 2006; received in revised form 21 September 2006; accepted 3 October 2006

Abstract

A laboratory-scale Activated Sludge System (ASS) was employed for the biodegradation of coke wastewater, which contains high concentrations of ammonium, thiocyanate, phenols and other organic compounds. The well-known kinetics models of Monod or Haldane are not very useful due to inhibition phenomena amongst the pollutants and also, they need the determination of a wide range of parameters to be introduced in the models. In this paper, a feed-forward neural network is outlined to obtain a satisfactory approach for estimating the effluent ammonium concentration of the treatment plant. ~~The methodology consists in performing a group of different sizes of the hidden layer and different subsets of input variables.~~

The developed model is useful to obtain simulations under different conditions of the influent stream, thus enabling the effluent ammonium concentration to be estimated. This neural network achieves better results than classical mathematical models for biological wastewater treatment as a result of the complex composition of the coke wastewater.

Keywords: Coke wastewater; Activated sludge; Neural network; Ammonium; Thiocyanate

1. Introduction

The wastewater used in this work originates from coke process in steelworks, which are present in most raw steel production facilities. Carbon requirements for iron smelting are obtained from the destructive distillation of coking coals at temperatures of between 900 °C and 1100 °C. When coal is heated in the absence of air, complex organic molecules within the coal break down to yield gases, liquid and solid organic compounds of lower molecular weight, and a nonvolatile carbonaceous residue which is known as coke.

The substances leaving the coke-ovens as liquids under ambient conditions are a flushing liquor consisting of free and fixed ammonium salts and other pollutants such as thiocyanate and cyanide; a tar containing several compounds that can be recovered, namely pyridine, tar acids, naphthalene, creosote oil and coal tar pitch and BTX aromatic hydrocarbon fractions; and an oil lighter than water that contains the compounds benzene, toluene, xylene and solvent naphthas.

Each of the three liquid streams undergoes further processing in the by-products section. The flushing liquor undergoes steam stripping, tar is recovered by removing the bottoms from settling tanks, and BTX's are extracted from the flushing liquor using liquid/liquid extraction. The resulting wastewater from these three processes makes up the coke wastewater. In this study, coke plant wastewater from the Arcelor Group steelworks in Avilés (Spain) was used.

Table 1 shows the average composition of coke wastewater, which was analysed daily over a 4-month period. As can be seen, the wastewater from the coke-making process contains considerable amounts of toxic compounds such as cyanide (31.8 mg/L), thiocyanate (363 mg/L) and phenols (207 mg/L).

* Corresponding author. Fax: +34 985 182 337.
E-mail address: emara@uniovi.es (E. Marañón).

1364-8152/$ - see front matter © 2006 Published by Elsevier Ltd.
doi:10.1016/j.envsoft.2006.10.001

Table 1
Composition of coke wastewater

Parameter	Coke wastewater (average value)
pH	8.1
Conductivity (mS/cm)	7.12
N-NH_4^+ (mg/L)	808
TKN (mg/L)	1040
COD (mg O_2/L)	1102
BOD_5 (mg O_2/L)	579
CN^- (mg/L)	31.8
PO_4^{3-} (mg/L)	0.54
TSS (mg/L)	32.0
VSS (mg/L)	23.0
NO_3^- (mg/L)	76.0
SCN^- (mg/L)	363
Phenols (mg/L)	207
Cl^- (mg/L)	1290
Fe (mg/L)	4.40
Zn (mg/L)	0.98

They also have high concentrations of ammonium, around 700 mg/L, and chlorides, above 1200 mg/L, but low concentrations of heavy metals and very low levels of phosphorus, around 0.5 mg/L. Hence, if the intention is to carry out a biological treatment, this nutrient will have to be added in the form of phosphates.

A laboratory-scale Activated Sludge System (ASS) was employed for the biodegradation of coke wastewater, with the following characteristics: (i) A wastewater homogenization tank, made of PVC with a total volume of 200 L, to which the nutrient needed for the biological process was added; and (ii) two aerobic reactors made of transparent PVC, with a total volume of 20 L. Oxygen was introduced in the reactor through orifices located at the bottom. A mechanical stirrer was employed to keep the liquor completely mixed. The temperature was kept constant at a value of 35 ± 0.5 °C by means of a heating element. (iii) Two settling tanks, also made of transparent PVC, with a total volume of 12 L, in order to return the settled sludge to aerobic reactors and thus keep the biomass concentration inside constant. Pumps were employed to feed the reactor and for recirculation.

In order to predict the behaviour of biological wastewater treatment plants, the Activated Sludge Model No. 1 (ASM1) of the International Association on Water Pollution Research and Control (IAWPRC) Task Group on Mathematical Modelling for Design and Operation of Biological Wastewater Treatment is frequently used (Henze et al., 1987). This simulation model distinguishes between heterotrophic and autotrophic biomass and different components of the wastewater, such as, for instance, the readily biodegradable substrate, the slowly biodegradable substrate and the soluble and particulate inert organic matter.

However, this model is not very useful to be applied to a type of wastewater with inhibition between pollutants like coke wastewater. For this reason, a feed-forward neuronal net was selected in this study to estimate the ammonium nitrogen concentration in the effluent. Similar applications using neural computing techniques can be found in Belanche et al. (1999), Capodaglio et al. (1991) and Steyer et al. (1997).

2. Process understanding and comprehension of the data

One of most complex tasks is the selection of variables to be used to model the system. An important working variable is the influent flow rate, which is correlated with the Hydraulic Residence Time (HRT) and the reactor volume.

Fig. 1 shows the influent flow rate as well as the concentrations of organic matter, expressed as Chemical Oxygen Demand (COD), ammonium nitrogen and thiocyanate. These compounds influence the removal of ammonium by

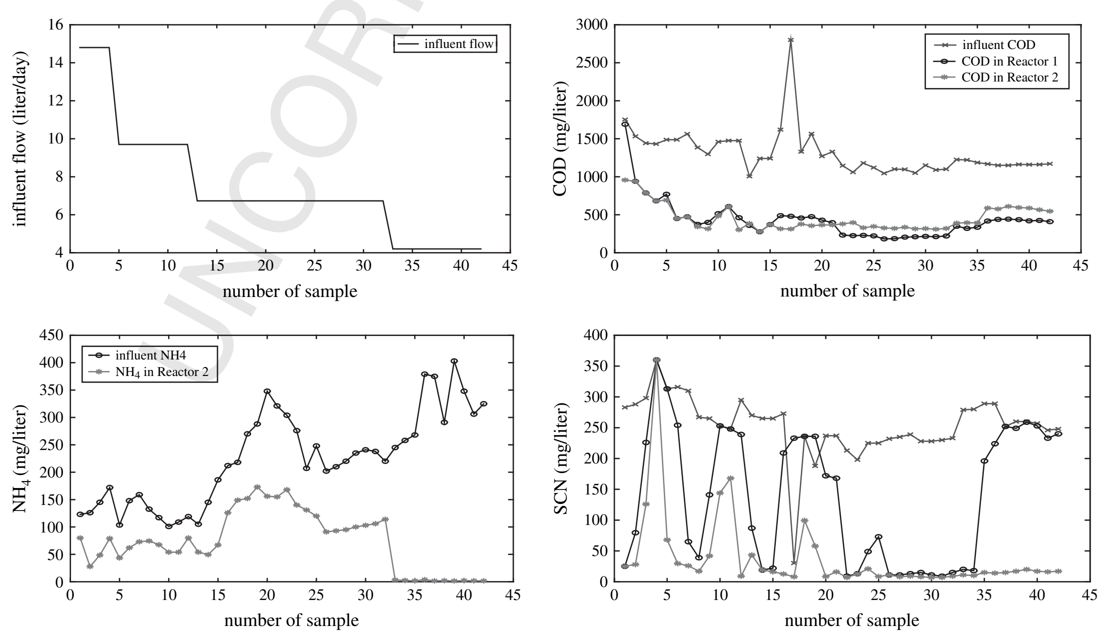

Fig. 1. Recorded values in laboratory pilot plant.

Please cite this article in press as: Machón, I. et al., Simulation of a coke wastewater nitrification process using a feed-forward neuronal net, Environmental Modelling & Software (2006), doi:10.1016/j.envsoft.2006.10.001

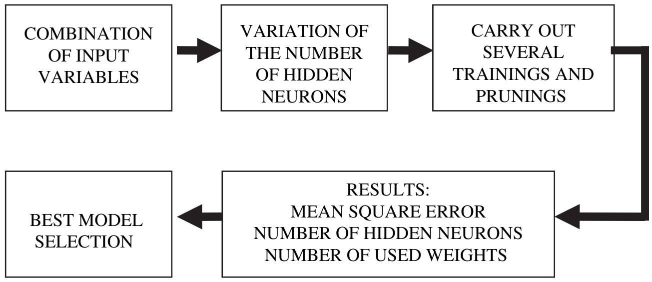

Fig. 2. Training and validation methodology.

nitrification. Nitrification takes place mainly in the second aerobic reactor, after the thiocyanate has been biodegraded in the first reactor. The pH in the first reactor was kept at 6.5 to favour thiocyanate biodegradation and at 8.4 in the second to favour nitrification. Furthermore, the solid retention time in the latter reactor (45 days) was sufficient for nitrification to take place. Accordingly, the influent ammonium concentration may be considered similar to the ammonium concentration in the first reactor.

When the COD increases, an increase in biological activity is required in order to remove the same amount of ammonium. Thiocyanate has a significant effect on nitrification, since its biodegradation produces ammonium. The biodegradation of thiocyanate occurs in both reactors and its concentration will be taken into account and integrated into the data set for training.

Finally, the ammonium concentration in the influent is the most important variable to consider, as it will influence the ammonium concentration in the effluent.

Taking all the above considerations into account, four variable sets will be analysed as input variables, namely

1. Influent flow, influent ammonium concentration, influent thiocyanate concentration and influent COD (Variable set No. 1).
2. Influent flow, influent ammonium concentration and influent COD (Variable set No. 2).
3. Influent flow and influent ammonium concentration (Variable set No. 3).
4. Influent flow, influent ammonium concentration and influent thiocyanate concentration (Variable set No. 4).

It is obvious that influent flow and influent ammonium concentration are key variables to estimate the effluent ammonium concentration. Moreover, COD and thiocyanate should be considered. For this reason, four data sets were formed by combination of these variables.

Several models of neural networks will be trained using these four variable sets in order to select the best model that minimizes the estimation error. Regarding the data acquisition, the measurements include 42 samples that were obtained approximately every 4–5 days. The data were divided into two sets of samples (training and testing) of the same size and the main characteristics of them were distributed between these two sets to achieve a better generalization.

3. Data pre-processing and training

The data are normalized to a zero mean and a unitary variance. This allows all the features to move in the same ranges and hence be treated by the neural network in the same way.

The architecture of the neural network is composed of a single hidden layer with hyperbolic tangent as the activation function and a single neuron with a linear activation function as the output layer. The activation function of the neurons of the hidden layer is a hyperbolic tangent. Thus, this type of function allows the network to learn nonlinear relationships. The activation function of the output layer is linear, thus enabling the network to take any value. This network topology can be used as a general approximator for any function that has a finite number of discontinuities whenever the hidden layer has a sufficient number of neurons and a nonlinear activation function (Hornik et al., 1989; Funahashi, 1989; Cybenko, 1989; Hartman et al., 1990).

The next step consists in training the neural network using the Levenberg–Marquardt algorithm (Levenberg, 1944; Marquardt, 1963; Moré, 1977). After training, the network is pruned, removing the weights that have the lowest saliences according to $H_{ij}\cdot w_{ij}^2/2$ (LeCun et al., 1990), where H_{ij} is the

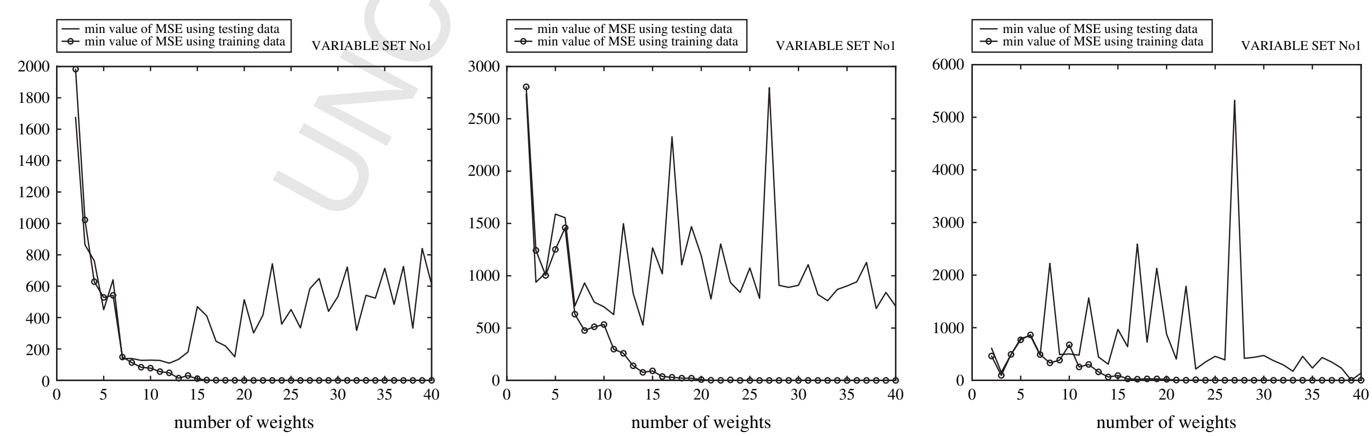

Fig. 3. Minimum value, mean value and standard deviation of mean square error of input variables: influent flow, influent SCN, influent COD and ammonium nitrogen concentration.

Please cite this article in press as: Machón, I. et al., Simulation of a coke wastewater nitrification process using a feed-forward neuronal net, Environmental Modelling & Software (2006), doi:10.1016/j.envsoft.2006.10.001

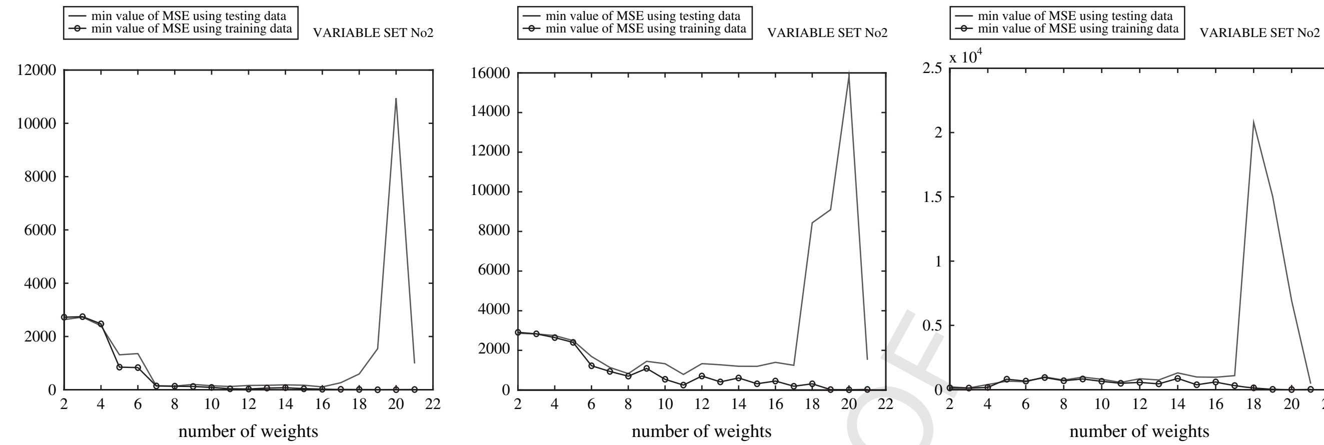

Fig. 4. Minimum value, mean value and standard deviation of mean square error of input variables: influent flow, influent COD and ammonium nitrogen concentration.

Hessian matrix and w_{ij} is the weight. At this stage, a data set different to that used for training is employed. The data set that is used is the testing data set.

4. Model selection

Discovering the input variables that optimize the approximation to the objective function is the first task to perform on the basis of the topology described above. The mean square error and the autocorrelation between the output variable and the error are useful to carry out this task (López et al., 2001).

Once the best combination of input variables has been selected, the number of hidden neurons must be determined. A low number of neurons does not provide enough parameters to train the neuronal network correctly. On the other hand, an excessive number of neurons leads to overtraining problems and its computational cost is higher. The training and pruning of the neuronal network was carried out using the toolbox named ''Neural Based Network Identification System'' developed by Helsinki Technical University (Norgaard et al., 2000, 2002).

Fig. 2 represents the methodology. The number of neurons in the hidden layer is increased gradually for each combination of input variables, making several tests with each of these (60 assays were carried out in this study). After pruning, the results are registered as the number of neurons in the hidden layer, the testing and training errors, the average and variance of these errors and the final number of used weights.

5. Results

The minimum value, mean value and standard deviation of the mean square error as a function of the number of used weights are calculated for each combination of variables. These are shown in Figs. 3–6. It can be seen that the higher the number of weights, the lower the training error. However, although the testing error decreases at the beginning, it rises later. Therefore, an optimum number of weights must be found.

Variable set No. 4 or the influent flow, the influent ammonium concentration and the influent thiocyanate concentration were chosen as input variables based on the minimum value,

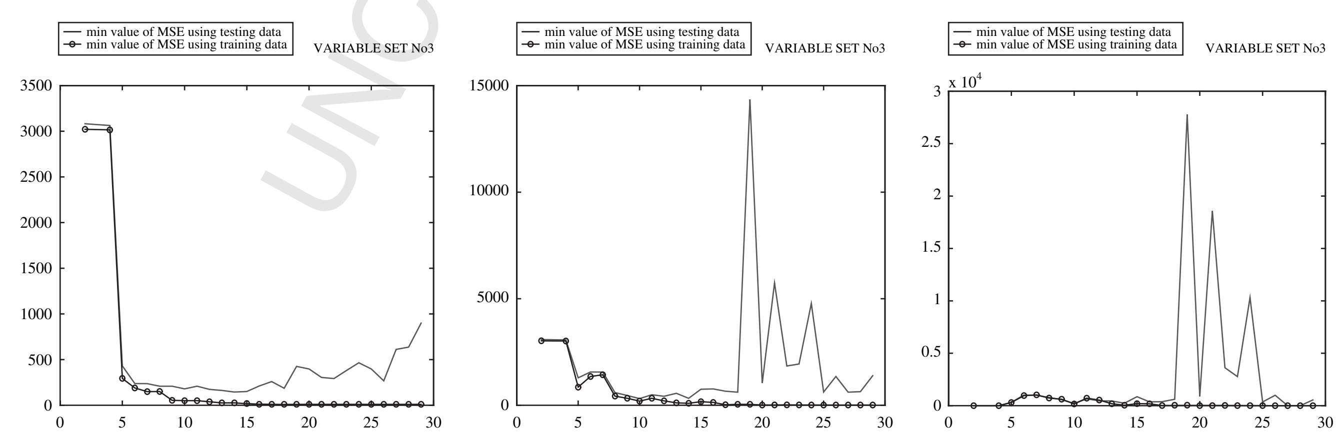

Fig. 5. Minimum value, average value and standard deviation of mean square error of input variables: influent flow and ammonium nitrogen concentration.

Please cite this article in press as: Machón, I. et al., Simulation of a coke wastewater nitrification process using a feed-forward neuronal net, Environmental Modelling & Software (2006), doi:10.1016/j.envsoft.2006.10.001

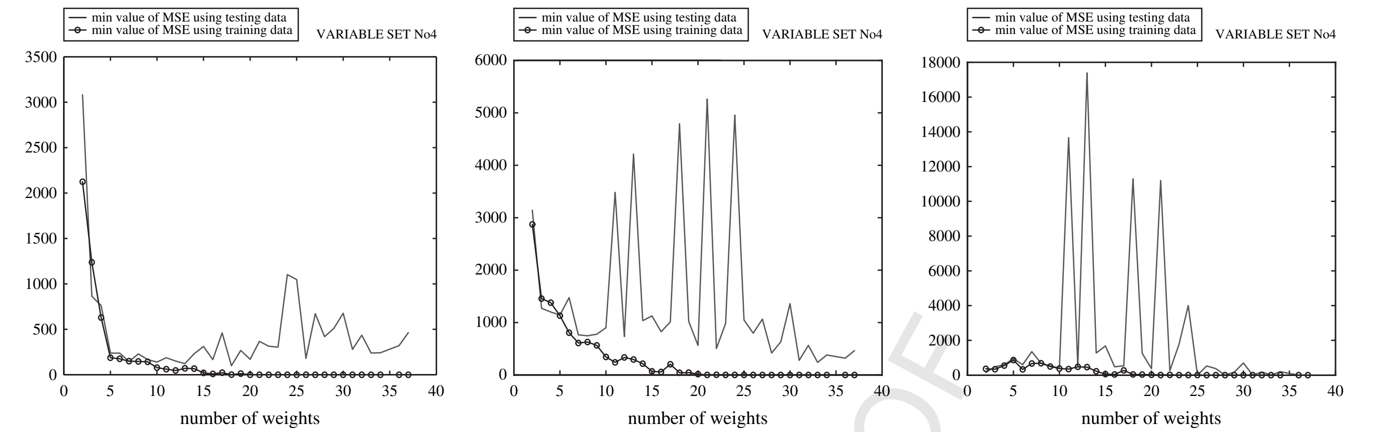

Fig. 6. Minimum value, mean value and typical deviation of mean square error of input variables: influent flow, influent SCN and ammonium nitrogen concentration.

the mean value and the standard deviation of the testing mean square error. The best results are obtained in a number of used weights equal to 18, which correspond to an original model of eight hidden neurons. Problems of local minima were detected in some models.

An iterative loop is established to search for the best model that minimizes the testing mean square error after pruning and considering the selected input variables and a number of hidden neurons equal to eight. Four of the eight neurons are not used in the best model after pruning. However, it is necessary to start training with a high enough number of neurons and then stop the procedure and remove the pruned weights.

Fig. 7 shows the real data, both training and testing data, corresponding to the effluent ammonium concentration and the values estimated by the neural network. The autocorrelation of the estimation error is good, tending quickly to zero and the distribution is also good, with most of the samples centered in the origin in accordance with Fig. 8.

6. Conclusions

A neural network model was developed to estimate the ammonium concentration in the effluent stream of a wastewater plant that undergoes biological treatment.

There are well-known mathematical models for biological treatment, such as the Activated Sludge Model (ASM). These can be formulated using kinetic dynamics (Haldane, Monod) and material balances to configure the particular structure of the plant. In this case, however, these models are not very useful because of the existence of inhibition between pollutants. Accordingly, the use of ANN's is recommended. Neural networks are widely used to estimate key parameters of physical processes.

In this paper, a feed-forward neural network is outlined to obtain a satisfactory approach to estimating the effluent ammonium concentration of the treatment plant. The methodology consists in performing a group of different sizes of the hidden layer and different subsets of input variables.

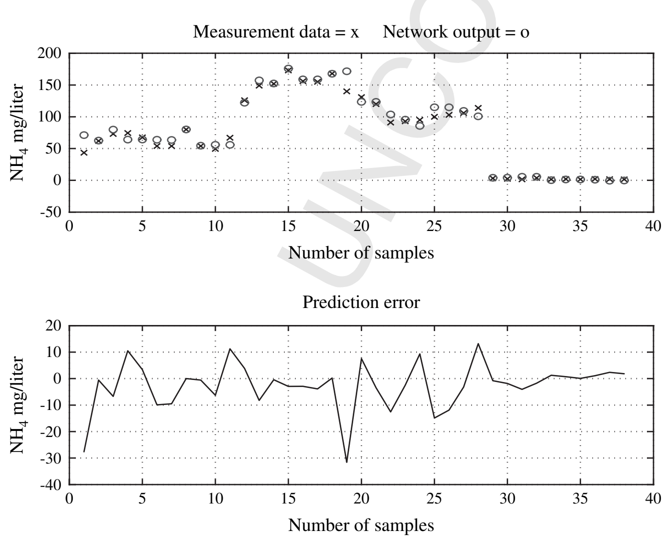

Fig. 7. Effluent ammonium nitrogen, real and simulated values.

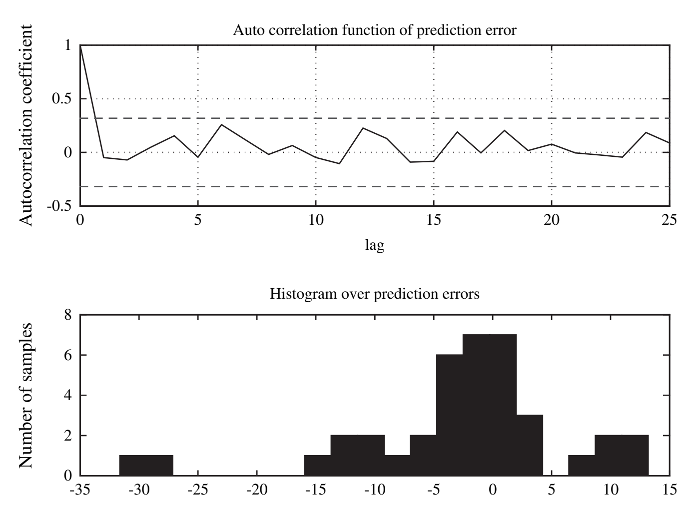

Fig. 8. Autocorrelation and error distribution.

Please cite this article in press as: Machón, I. et al., Simulation of a coke wastewater nitrification process using a feed-forward neuronal net, Environmental Modelling & Software (2006), doi:10.1016/j.envsoft.2006.10.001

The developed model is useful to obtain simulations under varying conditions of the influent stream, thus enabling the effluent ammonium concentration to be estimated. This neural network achieves better results than classical mathematical models for biological wastewater treatment as a result of the problematic composition of the coke wastewater.

As future work, the generalization ability and the accuracy may be improved training an ensemble of neural networks as can be seen in Torres et al. (2005) how the results are improved with this technique in other problems.

Acknowledgments

Our warmest thanks are expressed to the following for their financial support: The Commission of the European Communities, European Coal and Steel (ECSC), supporting the projects KNOWATER II “Implementation of a Knowledge Based System for Control of Steelworks Waste Water Treatment Plant” and BIOCONTROL “Advanced process control for biological water treatment plants in steelworks”, Ref. Nos. 7210-PR-234 and 7210-PR-235. ~~They~~ also wish to thank Paul Barnes for the English proof reading of this paper.

References

Belanche, Ll.A., Valdés, J.J., Comas, J., Roda, I.R., Poch, M., 1999. Towards a model of input–output behaviour of wastewater treatment plants using soft computing techniques. Environmental Modelling and Software 14 (5), 409–419.

Capodaglio, A.G., Jones, H.V., Novotny, V., Feng, X., 1991. Sludge bulking analysis and forecasting: application of system identification and artificial neural computing technologies. Water Research 25 (10), 1217–1224.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals, and Systems 2 (4), 303–314.

Funahashi, K.-I., 1989. On the approximate realization of continuous mappings by neural networks. Neural Networks 2 (3), ~~192–193~~.

Hartman, E.J., Keeler, J.D., Kowalski, J.M., 1990. Layered neural networks with Gaussian hidden units as universal approximations. Neural Computation 2 (2), 210–215.

Henze, M., Grady Jr., C.P.L., Gujer, W., Marais, G.v.R., Matsuo, T., 1987. Activated Sludge Model No. 1. IAWPRC Scientific and Technical Report No. 1. IAWQ, London, UK.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2 (5), 359–366.

LeCun, Y., Denker, J.S., Solla, S.A., 1990. Optimal brain damage. In: Touretzky, D.S. (Ed.), Advances in Neural Information Processing Systems, vol. 2. Morgan Kaufmann, San Mateo, CA.

Levenberg, K., 1944. A method for the solution of certain problems in least squares. Quarterly of Applied Mathematics 2, 164–168.

López, H., González, R., Machón, I., Ojea, G., Peregrina, S., González, J., de Abajo, N., 2001. Identification of melting and brightening section of a tinplate facility by means of neural networks. In: Proceedings of the European Control Conference, Porto.

Marquardt, D., 1963. An algorithm for least-squares estimation of nonlinear parameters. SIAM Journal on Applied Mathematics 11, 431–441.

Moré, J.J., 1977. The Levenberg–Marquardt Algorithm: Implementation and Theory. Numerical Analysis. In: Watson, G.A. (Ed.), Lecture Notes in Mathematics, vol. 630. Springer Verlag, Berlin.

Norgaard, M., Ravn, O., Poulsen, N.K., Hansen, L.K., 2000. Neural networks for modelling and control of dynamic systems. In: Doyle, A. (Ed.), Advanced Textbooks in Control and Signal Processing. Springer Verlag, London.

Norgaard, M., Ravn, O., Poulsen, N.K., 2002. NNSYSID-toolbox for system identification with neural networks. Mathematical and Computer Modelling of Dynamical Systems 8 (1), 1–20.

Steyer, J.-P., Rolland, D., Bouvier, J.-C., Moletta, R., 1997. Hybrid fuzzy neural network for diagnosis – application to the anaerobic treatment of wine distillery wastewater in a fluidized bed reactor. Water Science and Technology 36 (6–7), 209–217.

Torres, J., Hernández, C., Fernández, M., 2005. New results on ensembles of multilayer feedforward. artificial neural networks: formal models and their applications. Lecture Notes in Computer Science 3697, 139–144.

Please cite this article in press as: Machón, I. et al., Simulation of a coke wastewater nitrification process using a feed-forward neuronal net, Environmental Modelling & Software (2006), doi:10.1016/j.envsoft.2006.10.001