Parallel hybrid modelling methods for a full-scale cokes wastewater treatment plant

D.S. Lee*, P.A. Vanrolleghem* and J.M. Park**

* BIOMATH: Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Coupure Links 653, B-9000 Gent, Belgium

** Department of Chemical Engineering, School of Environmental Science and Engineering, POSTECH, San 31, Hyoja-dong, Pohang, Kyungbuk 790-784, Korea

Abstract Brid modelling methods are applied to model a full-scale cokes wastewater treatment plant. Within the hybrid model structure, a mechanistic model specifies the basic dynamics of the relevant process and non-parametric models compensate for the inaccuracy of the mechanistic model. Five different non-parametric models - feed-forward neural network, radial basis function network, linear partial least squares (PLS), quadratic PLS and neural network PLS - are incorporated in the hybrid model structure. For this application the hybrid model with neural network PLS as non-parametric model gives a better performance than the other non-parametric models in terms of the prediction accuracy, simplicity of use and interpretability.

Keywords Hybrid modelling; industrial wastewater treatment; mechanistic model; neural networks; partial least squares

Introduction

The activated sludge process is one of the most widespread wastewater treatment techniques for both municipal and industrial wastewater. Dynamic mathematical models enhance the understanding of the biological phenomena and provide the basis for design and operation of biological wastewater treatment systems. To date, the Activated Sludge Model No.1, or ASM1 (Henze *et al*., 1987), is accepted to be one of the most successful models for carbonaceous substrate and nitrogen removal processes in many applications. However, this model is limited in its ability to model industrial plants satisfactorily from a practical point of view. Every specific plant has its own process environmental conditions and process operations. It is not easy or worthwhile to spend too much time and effort to simulate peculiarities and non-idealities of an industrial process by using ASM1.

In recent years, hybrid modelling approaches have received considerable attention (Psichogios & Ungar, 1992; Thompson & Kramer, 1994; van Can *et al*., 1997; Lee *et al*., 2002). These approaches are potentially very efficient to obtain more accurate predictions of process dynamics by combining a mechanistic and a non-parametric model either in parallel or in serial configuration in such a way that the non-parametric model properly accounts for unknown and non-linear parts of the mechanistic model. There has been a great preference in artificial neural networks (ANN), especially feed-forward backpropagation neural networks (FBNN), as non-parametric model; however, the approach is applicable to other non-parametric models. The methods include radial basis function network (RBFN), linear partial least squares (PLS), quadratic PLS (QPLS) and neural network PLS (NNPLS).

In this study, the above mentioned five different non-parametric modelling methods were incorporated into a simplified mechanistic model in a parallel configuration. The methods were applied to a full-scale cokes wastewater treatment plant and were identified with the same data obtained from the plant. Models were tested on their prediction ability to extrapolate beyond the training data and compared based on their ease of model building, prediction accuracy and interpretability.

Methods and materials

Cokes wastewater treatment plant

The cokes wastewater treatment process (CWTP) at the steel-making company in Korea is a conventional activated sludge unit as shown in Fig. 1. It was designed for the removal of toxic organic pollutants from the cokes-making plant. The average composition of the cokes wastewater is shown in Table 1. Most of the chemical oxygen demand (COD) originates from phenol**,** which is a toxic inhibitory substrate but is also a carbon source for acclimatized microorganisms (Richards *et al*., 1989). In addition to phenol, cyanides and other toxic aromatic hydrocarbons such as cresol, indole, and toluene also contributes to the wastewater COD. Since a high concentration of nitrogen compounds was found inhibitory to biodegradation, pretreatment steps such as ammonia stripping were employed to render the wastewater more amendable to biodegradation. To alleviate the impact of high concentrations of deleterious substances on the biological treatment**,** an equalization tank was installed after the preliminary treatment stage and before the aeration tanks of the activated sludge process. The hydraulic retention time of the CWTP was approximately 2.7 days. Oxygen was introduced by submerged mechanical aerators. Concentrated sludge from the bottom of the clarifier was split into two streams**:** the first was recycled to the beginning of the first two aeration tanks and the other was treated in view of incineration of the waste sludge. The effluent from the settler was passed through chemical treatment units to remove hazardous heavy metal ions and to reduce the level of suspended solids and organic matter.

Operational data of five months (156 days) were collected at eight-hour intervals. All samples were analysed for mixed liquor suspended solids (MLSS), chemical oxygen demand (COD), suspended solids (SS), cyanide and phenol according to the Standard Methods (APHA, 1995). Dissolved oxygen concentration, pH, influent flow rate and recycle flow rate were also measured at each sampling time. This measuring campaign resulted in 466 operational data sets in total. The first 240 sets of data were used for training and the remaining 226 data sets for validation of the developed models.

Mechanistic Model

A mechanistic model for the CWTP was developed on the basis of ASM1. For the reactor system shown in Fig. 1, the mass balances for process variables in the reactor were established to predict the concentrations of effluent components. For practical reasons, minor mechanisms were neglected and only known major mechanisms were taken into account. Therefore, the following assumptions were made based on the process knowledge and understanding of the CWTP.

a) All reactors were assumed to be well-agitated and aerated continuous-flow reactors. The kinetics of microorganisms were assumed to be the same in all reactors.

b) Active biomass was divided into two types of organisms: heterotrophic (X_H) and cyanide-degrading organisms (X_{CN}).

c) Biodegradable carbonaceous material was classified as readily biodegradable substrate (S_S), slowly biodegradable substrate (X_S), and cyanide compounds (S_{CN}). Phenolic compounds in the

influent were assumed to be readily biodegradable organic matter (S_S). Inert particulate products (X_P) arising from biomass decay were also included.

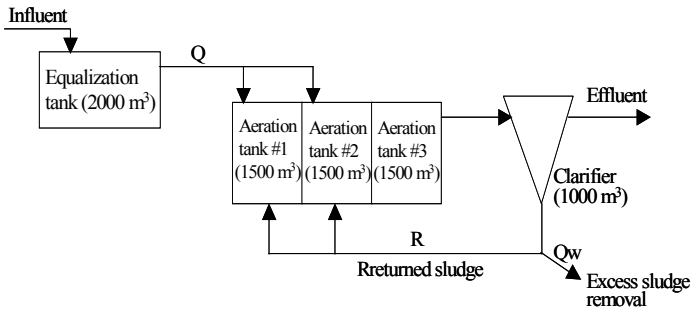


Figure 1 Schematic diagram of cokes wastewater treatment plant

Table 1 Average composition of cokes wastewater

Characteristics	Values
COD (mg/l)	1560
Phenol (mg-COD/l)	730
Cyanide (mg-COD/l)	9
m-cresol (mg/l)	65
Toluene (mg/l)	28
Indole (mg/l)	29

d) Nitrification reactions were neglected since no nitrate or nitrite was detected. There have been several reports that nitrification is inhibited by phenol, most of the poly-nuclear aromatic hydrocarbons, and cyanide, all of which were present as carbonaceous compounds in the cokes wastewater (Lee and Park, 1998).

e) Since the dissolved oxygen concentration was maintained at 1.5 mg O_2 /l**,** oxygen transfer was assumed not to be rate-limiting.

f) Cyanide compounds are toxic to heterotrophs. The shock loading of cyanides in wastewater caused a deterioration of the biological wastewater treatment processes. To describe cyanide inhibition upon the growth rate of heterotrophs, the following rate expression was used:

$$\mu = \hat{\mu}_H \left(\frac{S_S}{K_S + S_S}\right)\left(\frac{K_I}{K_I + S_{CN}}\right) \quad (1)$$

g) No reaction was assumed during the clarification and, therefore, the secondary settler was considered to be a simple separating point.

As a result, the mechanistic model consisted of eight components and five rate equations, as listed in Table 2.

The parameter values used in the model were initially based on literature values. For heterotrophs most default parameters provided by ASM1 were used. Kinetic parameters suggested by Gaudy *et al*. (1982) were used for cyanide-degrading organisms. From the results of sensitivity analysis only three parameters μ_H, μ_{CN} and Y_H were optimised using the simplex method in order to minimize the deviations between the simulation values and the corresponding operational data.

Parallel hybrid modelling methods

In the parallel hybrid model structure, non-parametric models described below are combined with the mechanistic model in a parallel configuration as shown in Figure 2. The non-parametric models were used to estimate the difference between the mechanistic model predictions and the corresponding operational data (*i.e*., the residuals). The key to success in developing parallel hybrid models lies in the information content of the residuals. If non-parametric models are trained to extract some useful information from the residuals, the accuracy of the model would be improved.

Since the mechanistic model was developed by making several assumptions and simplifications, we expected that the residuals would have dynamic information that was not contained in the mechanistic model. In addition, external disturbances such as composition variation, shock loadings of toxic compounds, and temperature variation were fed into the non-parametric model so that it could capture the effects of these disturbances. The inputs fed to both the non-parametric model and the mechanistic model include influent flow rate Q(t), sludge recirculation rate $Q_R(t)$ and the influent concentrations of COD(t) and cyanide(t). The effluent concentrations of COD(t-1), SS(t-1) and cyanide(t-1); MLSS(t-1)and pH(t-1) in the reactor; and the residuals e(t-1) of MLSS, SS, COD and cyanide were only fed to the non-parametric models. The non-parametric outputs were the residuals e(t) of MLSS, COD, SS and cyanide. All programs used in this work were implemented in MATLAB by using the Neural Network Toolbox (Demuth and Deale, 2001) and the PLS Toolbox (Wise and Gallagher, 2000).

Table 2 Simplified mechanistic model for the cokes wastewater treatment plant

Component →	*i*	1	2	3	4	5	6	7	8	Process rate ρ_j [$ML^{-3}T^{-1}$]
j	Process ↓	S_I	S_S	S_{CN}	X_I	X_S	X_H	X_{CN}	X_P	
1	Aerobic growth of heterotrophs		$-\frac{1}{Y_H}$				1			$\hat{\mu}_H\left(\frac{S_S}{K_S+S_S}\right)\left(\frac{K_I}{K_I+S_{CN}}\right)X_H$
2	Aerobic growth of cyanide-degrading microorganisms			$-\frac{1}{Y_{CN}}$				1		$\hat{\mu}_{CN}\left(\frac{S_{CN}}{K_{CN}+S_{CN}}\right)X_{CN}$
3	Decay of heterotrophs					$1-f_P$	-1		f_P	$b_H X_H$
4	Decay of cyanide-degrading microorganisms					$1-f_P$		-1	f_P	$b_{CN} X_{CN}$
5	Hydrolysis of entrapped organics		1			-1				$k_h\frac{X_S/X_H}{K_X+X_S/X_H}X_H$
Observed Conversion Rates [$ML^{-3}T^{-1}$]		$r_i=\sum_j \nu_{ij}\rho_j$								

[Stoichiometric papa meters]	[Kinetic parameters]
Heterotrophic yield: Y_H	Heterotrophic growth and decay: $\hat{\mu}_H$, K_S, K_I, b_H
Cyanide-degrading microorganisms' yield: Y_{CN}	Cyanide-degrading microorganisms' growth and decay: $\hat{\mu}_{CN}$, K_{CN}, b_{CN}
Fraction of biomass yielding particulate products: f_P	Hydrolysis: k_H, K_X

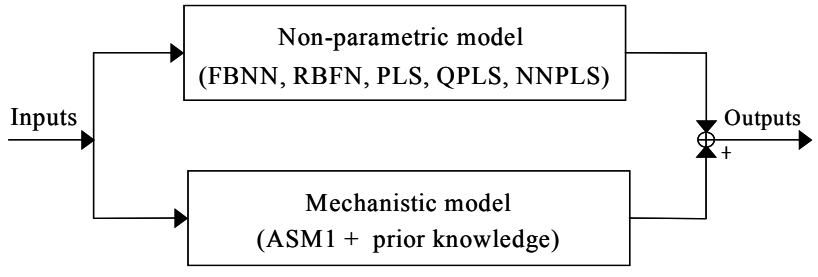


Figure 2 Schematic diagram of the parallel hybrid model

Feed-Forward Backpropagation Neural Network

Neural networks have been successfully applied to various biochemical processes (Montague *et al*., 1994; Zhang *et al*., 1994; Lee *et al*., 1999). They have a distinct ability to model non-linear dynamic systems without requiring a structural knowledge of the process to be modelled. Neural networks can map a set of input patterns onto a corresponding set of output patterns on the basis of historical data from any given system. However, neural network models have also been criticized for a lack of dependence upon physical relationships and a poor capacity for extrapolation.

A feed-forward backpropagation neural network (FBNN) was employed in this study. FBNNs have been successfully applied in modelling a wide range of non-linear systems, especially chemical/biological engineering processes (Baughman and Liu, 1995). The FBNN structure consists of one input layer, one hidden layer, and one output layer (Fig. 3a). Each layer can have a number of nodes (processing elements) which are connected linearly by weights to the nodes in the neighboring layers. The number of nodes in the input and output layers are predetermined by the number of input and output variables. In this study, the hidden layer has the hyperbolic tangent as activation function and the output layer the linear function (Fig.3b). Prior to training, all variables were scaled to the range –1 to 1. The Levenberg-Marquardt algorithm was used for the training process. The training process adjusts weights to minimize the error between the measured output and the output produced by the network. Through this adjustment, the neural network learns the input-output behaviour of the system. The optimal structure of the neural network was determined by varying the number of neurons in the hidden layer. By comparing the performance of network configurations in the recall and generalization process (Baughman and Liu, 1995), we found that the hidden layer with six neurons gave the best results.

Radial Basis Function Network

A radial basis function network (RBFN) is a network structure that employs local receptive fields to perform function mappings (Haykin, 1999). When an input vector is applied to the RBFN each neuron in the hidden layer will output a significant nonzero response according to how close the input vector is to each neuron's weight vector. The most frequently used radial basis function is the Gaussian activation function (Fig. 3c)

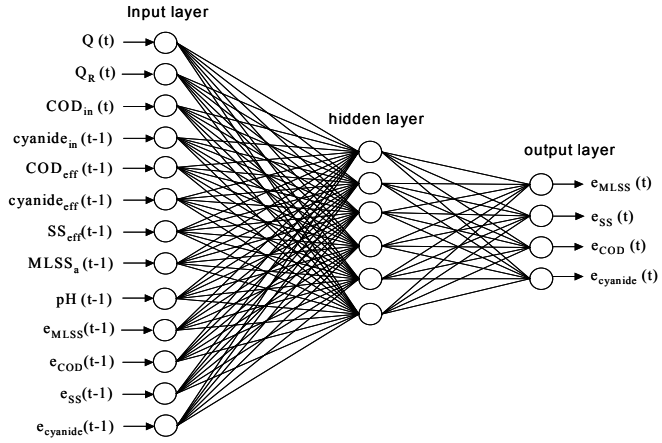
$$\Psi j(\mathbf{x}) = \exp\left(-\frac{\left\|\mathbf{x}-\mathbf{u}_j\right\|^2}{2\sigma_j^2}\right) \tag{2}$$

where $\mathbf{x}$ is an input vector, $\mathbf{u}_j$ is a weight vector and σ_j is the spread of the jth basis function. The output of the RBFN is the weighted average of the output associated with each hidden unit:

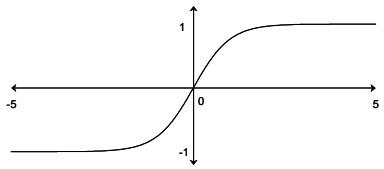
$$y_k(\mathbf{x}) = \frac{\sum_j w_{jk}\Psi_j(\mathbf{x})}{\sum_j \Psi_j(\mathbf{x})} \tag{3}$$

where $y_k(\mathbf{x})$ is the kth output and w_{jk} are the weights between the hidden and output layers. A gradient descent algorithm was used for training the network parameters. The RBFN was designed by creating neurons one at a time. Initially the hidden layer had no neurons. At each iteration the input vector with the greatest error was used to create a neuron. The training algorithm continued until the new network met the specified error goal or the maximum number of neurons was reached.

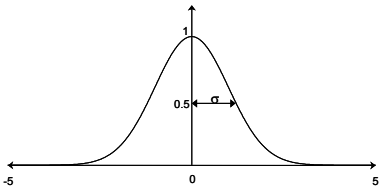
The spread parameter σ and the number of neurons in the hidden layer were optimised as 1.8 and 40, respectively, using response surface methodology.



(a) Architecture of one hidden layer neural network



(b) Hyperbolic tangent activation function



(c) Gaussian activation function

Figure 3 General architecture of neural network and activation functions

Partial Least Squares

The partial least squares (PLS) method is a linear multivariate method for relating the process variables **X** with responses **Y**. PLS can analyse data with strongly collinear, noisy and numerous variables in both **X** and **Y** (Wold *et al*., 2001).

PLS reduces the dimension of the predictor variables by extracting factors or latent variables which are correlated with **Y** while capturing a large amount of the variations in **X**. This means that PLS maximizes the covariance between matrices **X** and **Y**.

In PLS, the scaled matrices **X** and **Y** are decomposed into score vectors (**t** and **u**), loading vectors (**p** and **q**) and residual error matrices (**E** and **F**):

$$\mathbf{X} = \sum_{i=1}^{a} \mathbf{t}_i \mathbf{p}_i^T + \mathbf{E}$$
$$\mathbf{Y} = \sum_{i=1}^{a} \mathbf{u}_i \mathbf{q}_i^T + \mathbf{F} \qquad (4)$$

where a is the number of latent variables. In an inner relation the score vector **t** is linearly regressed against the score vector **u**.

$$\mathbf{u}_i = b_i \mathbf{t}_i + h_i \qquad (5)$$

where *b* is a regression coefficient which is determined by minimizing the residual *h*. There are several algorithms to calculate the PLS model parameters. In this work, the NIPALS (non-linear Iterative Partial Least Squares) algorithm was used with the exchange of scores (Geladi and Kowalski, 1986). It is crucial to determine the optimal number of latent variables and cross-validation is a practical and reliable way to test the predictive significance of each PLS component. Based on the cross-validation results six latent variables were included into the PLS model. It explained 76.4% of the variance of matrix **Y** and 87.7% of matrix **X**.

Non-linear Partial Least Squares

In order to capture non-linear structures between the predictor block and the responses, the PLS model can be extended to non-linear partial least squares models (Baffi *et al*., 2000). Major approaches have been to incorporate non-linear functions within the linear PLS framework. Especially, quadratic functions and neural networks have been used to identify the non-linear inner mapping between the input and the output latent variables. Wold *et al*. (1989) proposed the quadratic PLS (QPLS) method to make a polynomial fit for the PLS inner relation. QPLS works just like PLS and uses the NIPALS algorithm to calculate the latent variables. Once a pair of latent variables is calculated polynomial functions are used to model the functional relationship between the pair of latent variables. In our study, the degree of the polynomial used was two and the number of latent variables in the model was set to six on the basis of the cross-validation analysis.

Neural network PLS (NNPLS) is an integration of neural networks with PLS to model non-linear processes with input collinearity (Qin and McAvoy, 1992). The input and output variables are projected onto the latent space to remove collinearity and then each latent variable pair is mapped with a single input single output neural network. The neural network is trained to capture the non-linearity in the projected latent space. In this application, a FBNN with sigmoid functions was used to identify the non-linear inner regression models for each of the six latent variables.

Results and discussion

With the respective optimal models, the results of the validation stage enabled the different modelling approaches to be evaluated and compared. The performance of each model was evaluated in terms of the relative sum of squared error (RSSE) criterion. The RSSE performance index was defined as:

$$\mathrm{RSSE} = \frac{\sum \left\| y_{ij} - t_{ij} \right\|^2}{\sum \left\| t_{ij} \right\|^2} \tag{6}$$

where t_{ij} denotes the ith value of the output j and y_{ij} the associated prediction value. Table 3 shows the RSSE values of the training and validation data sets for six different modelling approaches. All models predict the dynamics of the wastewater treatment process with good accuracy compared to the mechanistic model. The parallel hybrid models even provide reasonable predictions during the cyanide shock loading period between days 83 and 87 as shown in Fig. 4 (Lee *et al*., 2002). This is a clear indication that the residuals contained sufficient relevant information about the dynamic behaviour of the system, but which was not considered in the mechanistic model. From the RSSE values, the parallel hybrid model with NNPLS as non-parametric model gave the best prediction performance.

Both the FBNN and RBF modelling structures provided a very good fit to the training data. However, the neural models showed a relatively poor predicting ability for the validating data sets compared with the non-linear PLS models. The RBF network was found to be superior to the FBNN model based on their validation performance and ease of optimisation. However, the RBF network required far more neurons in the hidden layer than the FBNN. This is because the Gaussian activation function in the RBF network only responds to relatively small regions of the input space, while the hyperbolic tangent function in the FBNN model can cover a large region of the input space. When the inputs vary over a wide range the required neurons in the RBF becomes high. The random initialisation of weights and bias in FBNN models gave slightly different simulation results during repeated training processes. On the contrary, initialisation of weights and bias in the RBF model is not a random process (Tetteh *et al*., 1996). Therefore it was possible to optimise the RBF network by a traditional factorial design for the spread parameter σ and the number of neurons in

the hidden layer. For both the FBNN and RBF models, it is a disadvantage that the weights and biases in the neural structures cannot be interpreted directly because of the relatively large number of parameters and the non-linear activation functions.

Table 3 Comparison of different parallel hybrid models

Non-parametric model	Structure	$RSSE_{recall}$	$RSSE_{validation}$
-	only mechanistic model	0.378	0.458
FBNN	13 - 6 - 4	0.134	0.404
RBF	13 - 40 – 4, sc= 1.8	0.110	0.361
PLS	latent variables= 6	0.237	0.429
QPLS	latent variables= 6, order of polynomial= 2	0.145	0.212
NNPLS	latent variables= 6, SISO FBNN	0.138	0.195

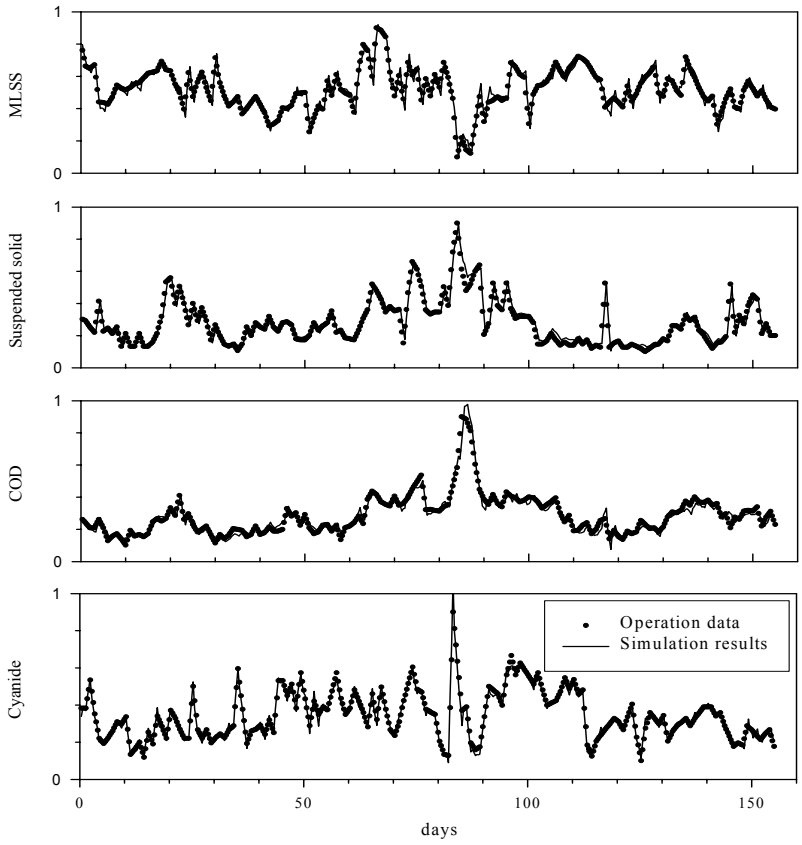


Figure 4 Simulation results using the parallel hybrid model with NNPLS as non-parametric model

It is clear from Table 3 that the PLS model could not adequately capture the dynamic behaviour of the process. The inability of the linear PLS approach to model the inner mapping between the input and output scores is evident. When non-linear regression methods were used for modelling the inner relations, the prediction capabilities were largely improved. The increased prediction performance of the non-linear PLS models can be explained by the fact that the CWTP, in this application, is an inherently non-linear dynamic system with time-varying reactions of the microorganisms and large variations in the incoming wastewater. QPLS gave a better prediction

performance than the PLS model. However, quadratic models are still linear in their parameters and do not guarantee a proper solution for mapping non-linear relationships between the input and output variables (Bro, 1995). The NNPLS method differs from the FBNN approach in that the data are not directly used to train the neural networks but are pre-processed by the PLS outer transform (Qin and McAvoy, 1992). The transform decomposes a multivariate regression problem into a number of univariate regressors. Therefore, the non-linear relationship between each pair of input and output scores can be modelled by a different network structure and over-fitting can be circumvented using the simple neural network models. NNPLS gave the best prediction performance among the non-parametric models in this study. This clearly indicates that a non-linear mapping exists between the input latent vectors and the response variable. These non-linear PLS models have also several interesting capabilities of visualizing high-dimensional data through the lower-dimensional projections formed by the score variables. In addition, it is straightforward to determine how much each of the inputs contributes to the variations in the output variables.

Conclusion

Five different hybrid modelling strategies were proposed to model a full-scale industrial wastewater treatment plant. First, a mechanistic model was developed based on ASM1 and specific process knowledge. Then the mechanistic model was combined with non-parametric models in parallel configuration. The non-parametric models evaluated include feed-forward backpropagation neural network (FBNN), radial basis function (RBF), partial least squares (PLS), quadratic PLS (QPLS) and neural network PLS (NNPLS). The neural network models FBNN and RBF could fit the training data very well, but could not adequately predict validation data compared with non-linear PLS models. The linear PLS model was not able to sufficiently capture the inherently non-linear process characteristics. Both non-linear PLS modelling techniques were much more successful. The NNPLS achieved the best prediction performance benefiting from the inclusion of a non-linear mapping between the input and output latent variables. In this study, the parallel hybrid model could improve the mechanistic model's predictions with available operational data. The proposed modelling strategy is a cost-effective and accurate tool that can be applied to biological wastewater treatment processes in the absence of reasonably accurate process models.

Acknowledgements

The work was financially supported by the Fund for Scientific Research – Flanders (F.W.O.) and the Ghent University Research Fund.

References

Baffi G., Martin E. B. and Morris A. J. (2000). Non-linear dynamic projection to latent structures modelling. *Chemom. Intell. Lab. Syst*., **52**, 5-22.

Baughman D. R. and Liu Y.A. (1995). *Neural networks in bioprocessing and chemical engineering*. Academic Press, San Diego.

Bro R. (1995). Algorithm for finding an interpretable simple neural network solution using PLS. *J. Chemon*., **9**, 423-430.

Demuth H. and Deale M. (2001). *Neural network toolbox for use with MATLAB*. The MathWorks, Massachusetts, USA.

Gaudy A. F. Jr., Gaudy, E. T., Feng Y. J. and Brueggemann G. (1982). Treatment of cyanide waste by the extended aeration process. *J. Water Pollut. Cont. Fed*., *54*, 153-164.

Geladi P. and Kowalski B. R. (1986). Partial least-squares regression: a tutorial. *Anal. Chim. Acta*, **185**, 1-17.

Haykin S. (1999). *Neural networks*. Prentice-Hall, New Jersey.

Henze M., Grady C. P. L. Jr, Gujer W., Marais, G. V. R., Matsuo T. (1987). Activated Sludge Model No. 1. IWA Scientific and technical report No. 1, London, UK.

Lee D. S., Jeon C.O., Park J. M. and Chang K. S. (2002). Hybrid neural network modelling of a full-scale industrial wastewater treatment process. *Bioeng. Biotech.*, **78**, 670-682.

Lee D. S. and Park J. M. (1999). Neural network modelling for on-line estimation of nutrient dynamics in a sequentially-operated batch reactor. *J. Biotechnol.*, **75**, 229-239.

Lee M. W. and Park J. M. (1998). Biological nitrogen removal from coke plant wastewater with external carbon addition. *Water. Environ. Res.*, **70**, 1090-1095.

Montague G. and Morris J. (1994). Neural-network contributions in biotechnology. *Trends Biotechnol.*, **12**, 312-324.

Psichogios D. C. and Ungar L. H. (1992). A hybrid neural network-first principles approach to process modeling. *AIChE J.*, **38**, 1499-1511.

Qin S. J. and McAvoy T. J. (1992). Non-linear PLS modelling using neural networks. *Comput. Chem. Eng.*, **23**, 395-411.

Richards D. J. and Shieh W. K. (1989). Anoxic-Oxic activated-sludge treatment of chanides and phenols. *Biotechnol. Bioeng.*, **33**, 32-28.

Standard Methods for the Examination of Water and Wastewater (1995). 19th edn, American Public Health Association/.American Water Works Association/Water Environment Federation, Washington DC, USA.

Tettch J., Metcalfe E. and Howells S. L. (1996). Optimisation of radial basis and backpropagation neural networks for modelling auto-ignition temperature by quantitative-structure property relationships. *Chemom. Intell. Lab. Syst.*, **32**, 177-191.

Thompson M., Kramer M. A. (1994). Modeling chemical processes using prior knowledge and neural networks. *AIChE J.*, **40**, 1328-1340.

van Can H. J. L., te Braake H. A. B., Hellinga C., Luyben K. C. A. M., Heijnen J. J. (1997). An efficient model development strategy for bioprocesses based on neural networks in macroscopic balances. *Biotechnol. Bioeng.*, **54**, 549-566.

Wise B. M. and Gallagher N. B. (2000). *PLS toolbox version 2.1 for use with MATLAB™*. Eigenvector Research, Washington DC.

Wise B. M., Holt B. R., Gallager N. B. and Lee S. (1995). A comparison of neural networks, non-linear biased regression and a genetic algorithm for dynamic model identification. *Chemom. Intell. Lab. Syst.*, **30**, 81-89.

Wold S., Sjöström M. and Eriksson L. (2001). PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.*, **58**, 109-130.

Wold S., Kettaneh-Wold N. and Skagerberg B. (1989). Non-linear PLS modelling. *Chemom. Intell. Lab. Syst.*, **7**, 53-65.

Zhang Q., Reid J. F., Litchfield J. B., Ren J. and Chang S.-W. (1994). A prototype neural network supervised control system for Bacillus thuringiensis fermentations. *Biotechnol. Bioeng.*, **43**, 483-489.