0273-2289/03/105-108/0437/$20.00

Simulation of Aerated Lagoon Using Artificial Neural Networks and Multivariate Regression Techniques

Karla Patricia Oliveira-Esquerre,*[,1] Aline C. da Costa,[1] Roy Edward Bruns,[2] and Milton Mori[1]

[1]DPQ/FEQ/UNICAMP, PO Box 6066, 13081-970 Campinas, SP, Brazil, E-mail: karla@feq.unicamp.br; and
[2]IQ/UNICAMP, PO Box 6154, 13083-970 Campinas, SP, Brazil

Abstract

The aim of this study was to develop an empirical model that provides accurate predictions of the biochemical oxygen demand of the output stream from the aerated lagoon at International Paper of Brazil, one of the major pulp and paper plants in Brazil. Predictive models were calculated from functional link neural networks (FLNNs), multiple linear regression, principal components regression, and partial least-squares regression (PLSR). Improvement in FLNN modeling capability was observed when the data were preprocessed using the PLSR technique. PLSR also proved to be a powerful linear regression technique for this problem, which presents operational data limitations.

Index Entries: Biochemical oxygen demand; functional link neural networks; partial least squares; principal components regression; multiple linear regression.

Introduction

Industrial and municipal wastewater are major sources of contamination of aquatic biota, accounting for several thousand types of chemicals released into the environment. Thus, the importance of implementing efficient monitoring and control techniques for wastewater treatment systems is well known. Operational control of biologic wastewater treatment plants is often complicated because of variations in raw wastewater compositions, strengths, and flow rates, owing to the changing and complex nature of the treatment process *(1)*. Moreover, the lack of suitable process variable measurements limits the effective control of effluent quality *(2,3)*.

*Author to whom all correspondence and reprint requests should be addressed.

The modeling traditionally used in bioprocesses is based on balance equations combined with rate equations for microbial growth, substrate consumption, and products formation. These methods have proven to be inefficient to represent the nonlinear, time-variant, and complex internal workings of the wastewater treatment process *(1,3–5)*.

Multivariate linear regression techniques may be used to approximate complex relationships over small regions of the predictor variables *(6)*. This approach is based on the assumption that the underlying nonlinear relationship can be locally approximated by a linear model. Partial least squares regression (PLSR), or projection to latent structures regression, has been shown to be a powerful linear regression technique for problems in which the data are noisy and highly correlated and in which there are only a limited number of observations *(6)*. PLSR provides the potential to summarize the underlying structures of the process as linear combinations of the original variables (latent variables [LVs]) and to relate the predictor and response variables by means of linear regression models among the LVs.

Recently, some studies using artificial neural networks (ANNs) in modeling biologic wastewater treatment processes have been published, providing an alternative approach *(1,3,4,7–12)*. In contrast to PLSR modeling, the main well-known drawbacks of neural network training are that it requires a large quantity of experimental data and its nonlinear approximation function can cause local minima problems. Moreover, the interpretation of such models is often difficult *(13)*.

There are no methods proposed to define the best structure to be used for a given case. A structure that has been little explored in bioprocesses is the functional link neural network (FLNN) *(13)*. In this network, a nonlinear functional transformation or expansion of the network inputs is initially performed and the resulting terms are combined linearly; that is, the estimation of network weights is linear. The structure obtained yields rapid convergence and good nonlinear approximation capability for a variety of functional approximation and inverse system modeling applications *(13,14)*.

Recent studies indicate that consideration of multivariate statistical principles in the neural network model building process may improve modeling performance *(15)*. Baffi et al. *(6)*, Kanjilal *(16)*, and Kompany-Zared *(17)* suggest the use of principal component (PC) variables, obtained by an orthogonal variable transformation, for pruning ANNs and improving their nonlinear mapping capabilities. PLSR LV can be used for the same purposes. Both techniques have been shown to be appropriate tools for extracting relevant information from correlated data. The use of principal component analysis (PCA) has been shown to be an efficient preprocessing technique for input data of ANNs *(9,17–19)*. Baffi et al. *(6)* and Holcomb and Morari *(19)* explored the combination of PLSR and ANN applied to nonlinear chemical processes.

The main aim of the present study was to develop an estimation model that provides accurate predictions of the biochemical oxygen demand

(BOD) of the output stream of an aerated lagoon at International Paper of Brazil, one of the major pulp and paper plants in Brazil. There is a 5-d delay in the determination of BOD, and when this is added to the 3-d hydraulic residence time, it is often too late to make proper adjustments in the wastewater treatment process.

In this work, the predictive models are calculated from FLNN and multivariate regression techniques (multiple linear regression [MLR], principal components regression [PCR] and PLSR). The potential improvement in FLNN modeling capability is also evaluated when PCs and LVs are used as the model's inputs.

The results show that FLNNs, MLR, nor PCR techniques are satisfactory when used individually for the aerated lagoon modeling and simulation. The best prediction performance is achieved when the data are preprocessed using PLSR before they are fed to an FLNN obtained by a first-order polynomial expansion and the logistic sigmoid activation as a transfer function. Similar results are obtained using a linear PLSR technique, showing that it can be appropriately used for modeling nonlinear systems.

Materials and Methods

The process wastewater was routed for primary treatment followed by biologic treatment in an aerated lagoon. To construct the predictive model, 10 variables of the aerated lagoon and 2 of the milling process were chosen using engineering judgment regarding which ones had an important effect on BOD prediction. Variations within sampling, including flow rate (FLOW), chemical oxygen demand (COD), suspended solids (S.S.), nitrate (N.N.) and ammonia (N.AM.) concentrations, conductivity (COND.), color (COLOR), pH, temperature (T.), rainfall (RAIN), and pulp (PULP) and paper (PAPER) production were evaluated. Their basic statistics are provided in Table 1.

The average hydraulic residence time in the aerated lagoon was used to establish a data input/output relationship. The original database, obtained from the plant control system and from the laboratory, has much unusable information that must be eliminated, so the 4-yr daily records available for investigation were reduced to 74 data points, each having 12 input variables and 1 output variable (BOD).

MLR, PCR, and PLSR

MLR models very often involve large numbers of predictor variables and their consequent collinearity or correlation provokes problems on regression. PCR and PLSR models perform an orthogonal transformation of data to form a set with reduced dimensionality; thus, there is no correlation between the transformed variables (PCs) used in the predictive regression, and the collinearity problem is solved. However, PCR is a two-step method and therefore there are the risks that useful predictive information ends up in discarded PCs and that some noise remains in the PCs used for regres-

Table 1
Basic statistical Descriptors for Selected Variables

Parameter	Average	Minimum	Maximum	SD	Skewness	Kurtosis
COD	624.19	297	881	99.84	–0.46	1.67
FLOW	61,103.81	35,156	73,693	5506.87	–1.74	6.15
S.S.	143.80	16	464	77.13	1.23	3.06
N.N.	1.55	0.568	5.99	0.93	2.47	8.52
N.AM.	2.70	0.03	9.2	1.52	1.43	3.88
COND.	1672.66	1036	3360	403.56	1.81	4.52
COLOR	442.69	111	911	159.65	0.44	1.06
pH	7.28	6.28	10.4	0.79	2.36	5.84
T.	45.85	29	50	3.20	–2.39	9.87
RAIN	4.17	0	40.7	9.47	2.49	5.35
PULP	894.24	0	1079.67	182.06	–3.26	12.15
PAPER	1060.31	685.3	1231.9	98.00	–1.66	4.19
BOD	83.23	20	155	26.44	0.35	0.30

sion *(20,21)*. Even though PLSR is closely related to the PCR method, PLSR is designed to maximize the prediction rather than fit the input data; that is, this method optimizes LVs in order to maximize the proportion of variance of their description of the prediction variables. Hense, the risk of losing predictive information is discarded or minimized. Mardia et al. *(22)* and Draper and Smith *(23)* provide detailed information on MLR and PCR. Geladi and Kowalski *(21)*, Wold et al. *(24)*, and Wold and Kowalski *(25)* provide details of the PLSR methodology.

FLNN

A conventional multilayer neural network contains one or more stages of neural processing between inputs and outputs. These stages, known as hidden layers, add to the complexity and cost of neural training *(13,26)*. When the number of inputs to the model and the number of records available for training becomes extremely large, the training procedure for this neural network architecture become increasingly more time-consuming *(27)*, whereas when the records available are too small, there is a problem of overfitting *(9)*.

The FLNN is a neural network with no hidden layers, trained using supervised learning. The input is "enhanced" by generating additional terms via some transformation rule, such as a polynomial expansion. The idea is to increase the dimensionality of the feature-space without requiring any additional information. These enhanced values are passed through to a summation node (the output), which transforms these weighted values via some nonlinear activation function. The use of the activation function is a modification of the structure of the FLNNs proposed by Henrique *(28)*. It increases the non-linear approximation ability of the network, while estimation of the parameters remains a linear problem *(26)*. In addition,

Henrique *(28)* proposed the use of an orthogonal estimator *(29)* to calculate the network weights and to eliminate nonlinear significant nodes during the training of the network *(30)*.

As can be seen, FLNN offers a remedy for time-consuming training and overfitting problems. It trains faster than multilayer networks because each interaction of the training procedure is linear and consequently takes less time. Its linear mapping requires a smaller data set for training and also avoids local minima problems.

Theoretical foundations of FLNN can be found in Costa et al. *(30)*, Hornik et al. *(31)*, and Pao *(32)*. The merits of the network output transformation proposed by Henrique *(28)* can be found in Costa et al. *(30,33)*, Harada et al. *(34)*, and Henriques et al. *(26)*.

Data Analysis by MLR, PCR, PLSR, and FLNN Modeling

Using a random selection method, 80% of all data records was assigned to multivariate model construction and FLNN training, while the remaining 20% was relegated to the validation set.

Stepwise regression was used to remove and add variables from the MLR model for the purpose of identifying a useful subset of predictors, reducing its dimensionality and collinearity *(23,35,36)*. $\alpha = 0.10$ (90% confidence level) and $\alpha = 0.15$ (85% confidence level) were considered for adding and removing variables, respectively.

Unfortunately, there is no universal, automatic criterion for the estimation of the number of PCs or LVs, and discrepancies among different criteria are common *(37)*. In the present work, the Todeschine *(38)* criterion was used. Todeschine *(38)* proposed a new correlation index (K) as a measure of quantity of correlation in the data set, as well as linear (KL) and nonlinear (KP) functions for estimating the maximum and minimum number of significant PCs to be used in the PCR model, respectively.

The optimal number of LVs is identified considering the increase in predicted accumulated variance (AV) that must be at least 2% for adding an LV *(39)*.

A polynomial expansion up to the third degree was performed on the FLNN inputs to generate nonlinear monomials and the orthogonal least-squares estimator proposed by Billings et al. *(29)* was used to calculate the network weights and to eliminate the monomials that are not significant in explaining the output variance. This reduces the size and complexity of the neural network and avoids overfitting the data *(26,30)*.

According to the modification of the structure of the FLNNs proposed by Henrique *(28)*, the network output is transformed by an invertible continuously differentiable nonlinear function. This activation function is chosen by trial and error to maximize the training performance of the FLNN. The three functions tested were

$$y = f(T) = \frac{1}{2} \log\left(\frac{1+T}{1-T}\right) \tag{1}$$

$$y = f(T) = \frac{1}{2}\log\left(\frac{T}{1-T}\right) \quad (2)$$

$$y = f(T) = \frac{1}{2}\log\left(\frac{0.5+T}{1-T}\right) \quad (3)$$

in which, Eq. 1 is a hyperbolic tangent sigmoid activation transfer function (Tansig), Eq. 2 is a logistic sigmoid activation transfer function (Logsig), and Eq. 3 is a centered logistic sigmoid activation transfer function (Logsic).

Finally, PCs and LVs were used as FLNN inputs and the prediction performance of the obtained models were compared.

p Values for the adjusted coefficient of multiple determination (R^2 adj.) and the normality test of the residual were calculated in order to verify the significance of the multivariate regression and neural network models, as well as to determine whether additional terms in the regression model would be useful. Both tests assumed a 95% confidence level. The *F* test *(40)* was used as an indicator of adequacy of model with different numbers of parameters. Root mean square error (RMSE) was calculated for modeling and validation results. Linear correlation index (R^2) values were calculated for the validation set.

In addition to regression graphs, three additional measures were used for the purpose of identifying outliers, or unusual observations that can have a significant influence on the regression: leverages, Cook's distance, and DFITS index *(35)*.

The same data sets were used for the training and validation of both neural network and multivariate regression models.

The Minitab, Statistica and Unscrambler computer programs were used for statistical analysis, data pretreatment; and MLR, PCR, and PLSR modeling. A Matlab computer program, developed by Henrique *(28)*, was used for the training and validation of the FLNN.

Results

The data were autoscaled to mean zero and unit variance prior to analysis. Table 2 provides the eigenvalues and the variance percentages (accounted for and cumulative) corresponding to the PCs for the predictor variable. PC modeling does not involve the predicted variable space. Table 2 also gives the variance percentages (accounted for and cumulative) corresponding to the LVs in both predictor and predicted variable space. Loadings were normalized to 1 and scores to their corresponding eigenvalues.

By taking a close look at the variance of the LVs and PCs, it can be verified that for any particular number of LVs or PCs, the LVs always describe more *predicted* variable variance and less *predictor* variable variance than the PCs.

Applying the Todeschine *(38)* criterion to the data matrix, we obtained $K = 0.3747$, which suggests a moderate degree of correlation, $KL = 8$, corresponding to 92.6% of AV and $KP = 5$ (75.7% AV) for PCA. This measure

Table 2
Statistical descriptors for PCs and LVs[a]

PC	Eigenvalue	EV (1) (%)	AV (1) (%)	LVs	EV (1) (%)	AV (1) (%)	EV (2) (%)	AV (2) (%)
1	3.6410	30.3	30.3	1	21.50	21.50	42.36	42.36
2	1.7805	14.8	45.1	2	20.01	41.51	12.99	55.35
3	1.3065	10.9	56.0	3	9.15	50.66	7.79	63.14
4	1.2502	10.4	66.4	4	7.51	58.17	4.43	67.57
5	1.1102	9.3	75.7	5	5.04	63.21	2.64	70.21
6	0.9454	7.9	83.6	6	7.08	70.29	0.49	70.70
7	0.5821	4.9	88.5	7	4.44	74.73	0.29	70.99
8	0.4977	4.1	92.6	8	6.22	80.95	0.13	71.12
9	0.3779	3.2	95.8	9	4.08	85.03	0.06	71.18
10	0.2579	2.1	97.9	10	8.16	93.19	0.01	71.19
11	0.1688	1.4	99.3	11	3.23	96.42	0	71.19
12	0.0819	0.7	100.0	12	3.58	100.00	0	71.19

[a]Explained variance (EV) and AV are given as predictor variable space (1) and predicted variable space (2).

reduced the number of PCs to five, which exhibited eigenvalues >1. This result agrees with the criterion suggested by Mardia et al. *(22)* and Morales et al. *(41)*, for which PCs whose eigenvalues are <1 ought to be excluded.

Considering the increase in the predicted AV, the optimal number of LVs was identified as five (which corresponds to 63.21 and 70.21% of the predictor and predicted AV, respectively). Only a 0.5% of increase in AV was reached with the inclusion of the sixth LV.

Figure 1 depicts the loading and score plots corresponding to the first two PCs (left) and PCs 3 and 4 (right). From the loading plots it can be deduced that the first PC (PC1) reflects an inverse correlation between COND. and FLOW, and T. and PULP (loadings –0.42, 0.41, 0.43, and 0.47, respectively). The second PC (PC2) is associated with COD and suspended solids S.S. (loadings –0.51 and –0.58, respectively). The third PC (PC3) reflects a negative correlation between N.N. and COLOR (loadings –0.56 and 0.49, respectively). The fourth PC (PC4) is related to COLOR, RAIN), pH, N.N., and N.AM. data (loadings 0.48, 0.38, 0.38, –0.43, and –0.40, respectively). As can be seen, PC1, PC3, and PC4 are difficult to interpret. PC2 may be accounted for by a microbiologic factor, and the fifth PC (PC5) is exclusively related to PULP (loading 0.78) (not shown). There is a relatively spread-out distribution among the PCs, which is quite common when working with environmental data *(41)*.

As expected, there is no indication that the PCs are correlated in the score plots. However, some samples exhibit more anomalous scores than the rest, so they can be indicated as possible outliers.

Consistent with the stepwise results, the COD, COND., T. and PULP, and PAPER parameters appear to be the most significant in the first five

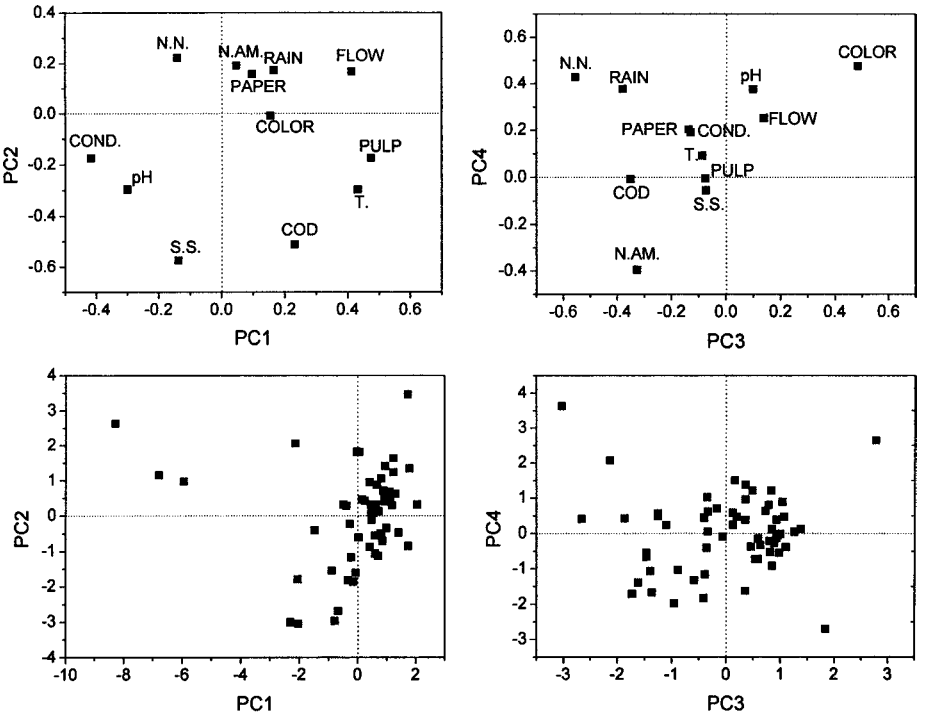

Fig. 1. Loading (top) and score plots (bottom), corresponding to first four PCs.

LVs. As expected, since the COD is the parameter most correlated with BOD, the first LV is exclusively related to the COD (loading 0.70) and COD is the first parameter included in the MLR by the stepwise method. The multivariate regression models were then constructed and their prediction performance results are given in Table 3. As expected, MLR, PCR, and PLSR models present the same performance results when there is no parameter exclusion. The p values for the adjusted coefficient of multiple determination (R^2 adj.) are zero for all models, indicating their high statistical significance. In spite of this, the p values for the normality test of the residuals and the F test for model adequacy analysis indicate that predictive information has been lost by exclusion of the last seven PCs. It had not occurred when the stepwise method was used for identifying the most significant PCs for the predictive model (eighth and ninth have been excluded). Only the stepwise-MLR model residuals did not pass in the normality test. From the modeling and validation statistical results, it can be seen that best prediction performance was achieved using the shortest PLSRR model.

The FLNN models were then generated. The number of monomials generated is a function of the number of inputs and the degree of the polynomial expansion, in this case 12 and 3, respectively. Then, the number of monomials generated is 455. Using the orthogonal least-squares estimator,

Table 3
Statistical Results of Multivariate Regression Models

		Modeling			Validation	
Model	R^2 adj.	RMSE	p Value[a]	F test[b]	R^2	RMSE
MLR	0.637	0.063	0.108	0.98	0.619	0.070
Stepwise-MLR	0.636	0.068	0.003		0.656	0.069
PCR	0.637	0.063	0.108	26.61[c]	0.619	0.070
PCR (1–5 PCs)	0.320	0.092	0.786		0.379	0.095
PCR (1–7, 10–12 PCs)	0.638	0.064	0.181	7.60[d]	0.615	0.070
PLSR	0.637	0.063	0.108	0.22	0.619	0.070
PLSR (1–5 LV)	0.674	0.064	0.250		0.666	0.067

[a]p Value for the normality test of the residuals.
[b]Test indicated by Barros et al. *(40)*. $F_{0.05, 7, 46} = 2.22$
[c] Between PCR and PCR (1–5PC) models and [d] between PCR and PCR (1–7, 10–12 PC) models.

Table 4
Statistical Results of FLNN, PCA-FLNN, and PLSR-FLNN Models

		Modeling			Validation	
Model	R^2 adj.	RMSE	p Value[a]	F test[b]	R^2	RMSE
FLNN	0.647	0.062	0.181	0.76	0.555	0.096
Stepwise-FLNN	0.659	0.066	0.003		0.585	0.098
PCA-FLNN	0.647	0.062	0.181	29.34[c]	0.602	0.083
PCA-FLNN (1–5 PCs)	0.198	0.094	0.575		0.350	0.096
PCA-FLNN (1–7, 10–12 PCs)	0.629	0.064	0.679	1.19[d]	0.604	0.086
PLSR-FLNN	0.647	0.062	0.181	0.31	0.648	0.076
PLSR-FLNN (1–5 LVs)	0.679	0.063	0.297		0.717	0.069

[a]p Value for the normality test of the residuals.
[b]Test indicated by Barros et al. *(40)*. $F_{0.05, 7, 46} = 2.22$.
[c]Between PCA-FLNN and PCA-FLNN (1–5 PC) models and $F_{0.05, 2, 46} = 3.26$.
[d]Between PCA-FLNN and PCA-FLNN (1–7, 10–12 PC) models.

nonsignificant monomials were eliminated and the weights are estimated. The best neural network prediction results are given in Table 4.

By comparing the FLNN approach with the approaches proposed herein, PCA-FLNN and PLSR-FLNN, it can be seen that the best prediction performance was achieved when the data were preprocessed using PLSR before they were fed to an FLNN using a first-order polynomial expansion and the logsig activation function. Considering the validation set results, PLSR-FLNN modeling performance was significantly improved when only the most significant LVs were used as input. However, this pruning

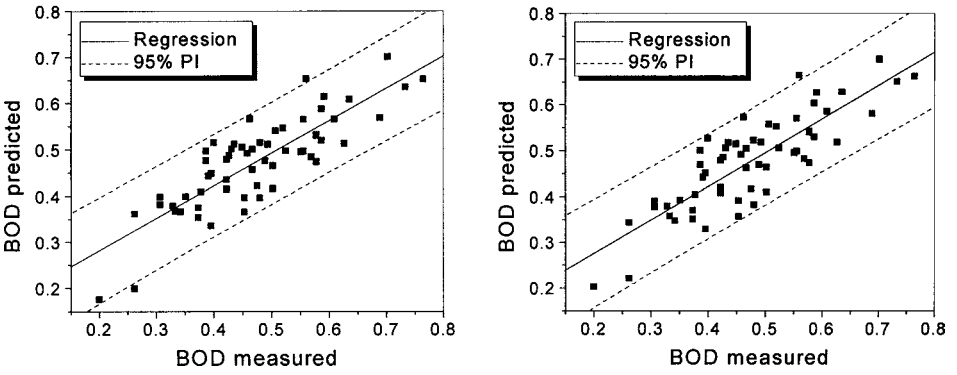

Fig. 2. Relation between predicted vs measured BOD (solid lines). Upper and lower dashed lines indicate the 95% prediction estimation interval (according to the PLSR and PLSR-FLNN models, without the last seven PCs).

does not help FLNN and PCA-FLNN prediction results for the validation data set.

Comparison of multivariate regression and neural network results, reveals that PLSR and PLSR-FLNN models gave similar prediction results when the last seven LVs were excluded. It really shows that the linear PLSR technique can be used to approximate complex relationships over the regions of the predictor variables available for the considered process. However, for the validation set, PLSR-FLNN gave slightly better results.

Statistical results also show that the best modeling structures—i.e., PLSR and PLSR-FLNN without the last seven PCs—although still far from the perfect, do capture prediction information and are able of estimating outputs within a 95% prediction band (*see* Fig. 2). This is particularly important when one takes into account the complexity of the wastewater treatment system and the large quantity of missing data in the modeling set.

Figure 3 presents a graphic representation of the measured and predicted BOD data for the modeling and validation data sets using PLSR and PLSR-FLNN models without the last seven PCs. The modeling features were well reproduced by both modeling structures, whereas the validation features were slightly better reproduced by the PLSR-FLNN model.

Multivariate regression and the neural network have also provided robust models. Their model parameters do not change very much when the samples identified as possible outliers were taken from the training data set.

Discussion

In recent years, ANNs have become extremely popular for prediction and forecasting in a number of areas, including water resources, bioprocesses, and environmental science. The FLNN appeared as a good

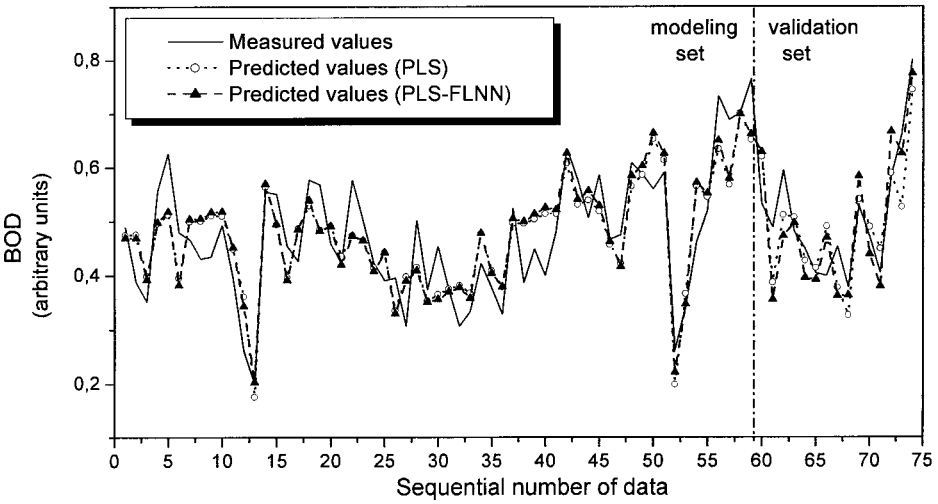

Fig. 3. Measured and predicted BOD, using the PLSR and PLSR-FLNN models (without last seven PCs).

option to remedy the time-consuming training and overfitting problems of the classic feedforward back-propagation networks.

In the present work, some modifications were used in the functional link network structures, such as the use of an invertible activation function, which has been proved to increase the nonlinear approximation capability.

As in PCR and MLR models, the simple structure of FLNNs led to an unsatisfactory performance for the simulation and prediction of BOD for the data set treated here. Then, in addition to the method used to eliminate nonsignificant monomials during the network training, PCA and PLSR were used as input pretreatment techniques for pruning the FLNN structure. The combined use of these techniques and FLNNs provided prediction results that have statistical parameters significantly superior to those obtained using the simple FLNN structure.

Best prediction results were given using the PLSR and FLNN structure obtained by using a first-order polynomial expansion and the logistic sigmoid activation as a transfer function, when the last seven latent variables were excluded. The PLSR technique helped FLNN mapping by its orthogonal transformation of variables and reduction of the system's dimensionality.

Our work also shows the advantage of linear PLSR in its ability to represent highly nonlinear relationships, even for a system that presents operational data limitations (e.g., imprecision associated with measured variables, a limited range of variables, a large number of missing values). In spite of the nonlinear nature of the process, linear PLSR gave similar prediction results for the modeling data set and slightly inferior ones for thevalidation set when compared with the PLSR-FLNN approach.

Acknowledgments

We are grateful to professor Dr. Fernando Von Zuben, Vitaly F. Rodríguez Esquerre, Roberto C. Colacioppo, and Celeste M. Díaz Cónsul for the many stimulating discussions held during the development of this work; and to International Paper of Brazil for providing the data set. We also thank Fundação de Amparo à Pesquisa do Estado de São Paulo (Proc. N. 99/10257-0) for their financial support.

Nomenclature

BOD	=	outlet wastewater BOD (mg/L)
COD	=	inlet wastewater COD (mg/L)
COLOR	=	color (ppm or mg/L)
COND.	=	conductivity (μS/cm at 20°C)
f	=	transfer function proposed by Henrique *(28)*
FLOW	=	inlet flow rate (m^3/d)
h	=	polynomial expansion
N	=	number of monomials
N.AM.	=	inlet ammonia concentration (mg/L)
N.N.	=	inlet nitrate concentration (mg/L)
PAPER	=	paper production (t/d)
PULP	=	pulp production (t/d)
RAIN	=	rainfall (mL/d)
S.S.	=	inlet suspended solids (mg/L)
T.	=	temperature (°C)
$\mathbf{x}$	=	input matrix
y	=	predicted output
w_{ij}	=	neural network weights of i input and j monomial

References

1. Hamoda, M. F., Al-Ghusain, I. A., and Hassan, A. H. (1999), *Water Sci. Technol.* **40,** 55–65.
2. Harremoës, P., Capodaglio, A. G., Hellstrom, B. G., Henze, M., Jensen, K. N., Lynggaaard-Jensen, A., Otterpohl, R., and Soeborg, H. (1993), *Water Sci. Technol.* **27,** 71–115.
3. Lee, D. S. and Park, J. M. (1999), *J. Biotechnol.* **75,** 229–239.
4. Cote, M., Grandijean, B. P. A., Lessard, P., and Yhibault, J. (1995), *Water Res.* **29,** 995–1004.
5. Steyer, J. P., Rolland, D., Bouvier, J. C., and Moletta, R. (1997), *Water Sci. Technol.* **36,** 209–217.
6. Baffi, G., Martin, E. B., and Morris, A. J. (1999), *Comp. Chem. Eng.* **23,** 1293–1307.
7. Gontarski, C. A., Rodrigues, P. R., Mori, M., and Prenem, L. F. (2000), *Comp. Chem. Eng.* **24,** 1719–1723.
8. Häck, M. and Köhne, M. (1996), *Water Sci. Technol.* **33,** 101–115.
9. Oliveira-Esquerre, K. P., Mori, M., and Bruns, R. E. (2002), *Braz. J. Chem. Eng.* **19,** 365–370.
10. Pu, H., Hung, Y. (1995), *Environ. Manage. Health* **6,** 16–27.
11. Wilcox, S. J., Hawkes, D. L., Hawkes, F. R., and Guwy, A. J. (1995), *Water Res.* **29,** 1465–1470.

12. Zhao, H., Hao, O. I., Fellow, A. S. C. E., McAvoy, T. J., and Chang, C. H. (1997), *J. Environ. Eng.* **123,** 311–319.
13. Chen, S. and Billings, S. A. (1992), *Int. J. Control* **56,** 319–346.
14. Alkulaibi, A. and Soraghan, J. J. (1997), *Signal Processing* **62,** 101–109.
15. Maier, H. R. and Dandy, G. C. (2000), *Environ. Modelling Software* **15,** 101–124.
16. Kanjilal, P. P. (1995), *IEEE Trans. Neural Networks* **6,** 1061–1070.
17. Kompany-Zared, M. (1999), *Talanta* **48,** 283–292.
18. Cancilla, D. A. and Fang, X. (1996), *J. Great Lagoons Res.* 22, 241–253.
19. Holcomb, T. R. and Morari, M. (1992), *Comp. Chem. Eng.* 16, 393–411.
20. Despagne, F. and Massart, D. L. (1998), *Analyst* 123, 157–178.
21. Geladi, P. and Kowalski, B. R. (1986), Anal. Chim. Acta 185, 1–17.
22. Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979), *Multivariate Analysis*, Academic, London, UK.
23. Draper, N. R. and Smith, H. (1981), *Applied Regression Analysis*, 2nd ed., Wiley, New York, NY.
24. Wold, S., Martens, H., and Russwurm, H. (1983), *Food Research and Data Analysis*, Applied Science Publishers, London, UK.
25. Wold, S. and Kowalski, B. (1984), *Chemometrics: Mathematics and Statistics in Chemistry*, Reidel, Dordrecht, The Netherlands.
26. Henriques, A. W. S., Costa, A. C., Alves, T. L. M, and Lima, E. L. (1999), *Braz. J. Chem. Eng.* **16,** 171–177.
27. Cass, R. and Radl, B. (1996), *Control Eng. Pract.* **4,** 1579–1584.
28. Henrique, H. M. (1999), PhD thesis, PEQ/COPPE/UFRJ, Rio de Janeiro, RJ, Brazil.
29. Billings, S. A., Chen, S., and Korenberg, M. J. (1989), *Int. J. Control* **49,** 2157–2189.
30. Costa, A. C., Henriques, A. S. W., Alves, T. L. M., Maciel Filho, R., Lima, E. L. (1999), *Braz. J. Chem. Eng.* **16,** 53–63.
31. Hornik, K., Stinchcombe, M., and White, H. (1989), *Neural Networks* **2,** 359–366.
32. Pao, Y. H. (1989), *Adaptative Pattern Recognition and Neural Networks*, Addison-Wesley, Reading, MA.
33. Costa, A. C., Alves, T. L. M., Henriques, A. W. S., Maciel Filho, R., and Lima, E. L. (1998), *Comp. Chem. Eng.* **22(Suppl.),** S859–S862.
34. Harada, L. H., da Costa, A. C., and Maciel Filho, R. (2002), *Appl. Biochem. Biotechnol.*, **98–100,** 1009–1023.
35. Montgomery, D. C., Peck, E. A. (1992), *Introduction to Linear Regression Analysis*, Wiley, New York, NY.
36. Sjöstrom, M. and Wold, S. (1983), *Anal. Chim. Acta* **150,** 61–70.
37. Ortiz-Estarelles, O., Martín-Biosca, Y., Medina-Hernández, M. J., Sagrado, S., and Bonet-Domingo, E. (2001), *Chem. Intel. Lab. Syst.* **56,** 93–103.
38. Todeschine, R. (1997), *Anal. Chim. Acta* **348,** 419–430.
39. Eigenvector Research (1998) MATLAB PLS Toolbox help. Eigenvector Research Inc., Manson, WA
40. Barros, B. N., Scarminio, I. S., and Bruns, R. E. (1995), *Planejamento e Otimização de Experimentos*, UNICAMP Press, Campinas, SP, Brazil.
41. Morales, M. M., Martí, P., Llopis, A., Campos, L., and Sagrado, S. (1999), *Anal. Chim. Acta* **394,** 109–117.