Desalination and Water Treatment
www.deswater.com

doi: 10.1080/19443994.2016.1191777

57 (2016) 27969–27977
December

Taylor & Francis
Taylor & Francis Group

Prediction of organic matter removal from pulp and paper mill wastewater using an artificial neural network

Jácina T.G. Morais[a], Karla P.S.O.R. Esquerre[b], Asher Kiperstok[c], Luciano M. Queiroz[c,*]

[a]*Department of Chemical Engineering, Federal University of the Semi-Arid Region, Caraúbas, Rio Grande do Norte 59780-000, Brazil, Tel./Fax: +55 84 33272676; emails: tabitagurgel@yahoo.com.br, verlandia.peixoto@ufersa.edu.br (J.T.G. Morais)*
[b]*Department of Chemical Engineering, Federal University of Bahia, Salvador, Bahia 40210-630, Brazil, Tel./Fax: +55 71 32839800; emails: karla.esquerre@gmail.com, cei@ufba.br (K.P.S.O.R. Esquerre)*
[c]*Department of Environmental Engineering, Federal University of Bahia, Salvador, Bahia 40210-630, Brazil, Tel./Fax: +55 71 32839785; email: asher@ufba.br (A. Kiperstok), Tel. +55 71 32839796; Fax: +55 71 32839454; emails: lmqueiroz@ufba.br, maasa@ufba.br (L.M. Queiroz)*

Received 27 October 2015; Accepted 15 May 2016

ABSTRACT

The Multilayer Perceptron Model was developed for predicting organic matter removal from pulp and paper mill wastewater. The original database covered a period of 1,427 consecutive days and contained the most frequently measured parameters. Three models were constructed by applying the technique of Principal Component Analysis, which selected principal components, discarded original variables and excluded possible outliers. The data were randomized and divided into training, validation and testing sets. The training algorithm was the Levenberg–Marquardt type, which is an adaptation of the back-propagation algorithm. The learning rate was 0.05, and the evaluation criteria used were the mean square error and the linear correlation coefficient. A marked difference was observed in the predictive performance when the organic matter load was used as an input. The model M4, which was built by discarding the two variables pH and EC, proved to be the most suitable and the simplest model obtained. However, choosing the best model cannot be done arbitrarily. It will be necessary to use various statistical parameters and perform comparisons of models with different sizes and structures in order to select the best model.

Keywords: Artificial neural network; Organic matter; Multilayer perceptron model; Pulp and paper mill wastewater

1. Introduction

Safer operation and control of industrial processes can be achieved by developing a modeling tool to predict plant performance, which can be based on past observations of certain key product quality parameters. Performance assessment and monitoring of biological wastewater treatment processes are usually made by collecting samples and conducting physical–chemical analysis on a daily or weekly basis, which

*Corresponding author.

Presented at the International Conference on Industrial Waste and Wastewater Treatment and Valorization (IWWTV 2015) May 21–23, 2015, Athens, Greece

leads to an increase in the overall cost of the process. Moreover, the numerical modeling to quantify the efficiency of contaminants removal, e.g. organic matter, is based on models with kinetic constants usually obtained from studies of isolated cultures of micro-organisms fed with specific substrates under the laboratory scale. Microbial diversity and variability of organic substrates supplied to the micro-organisms, which are associated with the variation in operating conditions in industrial processes, may limit the use of specific kinetic models for predicting the performance of wastewater treatment systems. Therefore, utilizing a series of monitoring data, predictive statistical tools are attractive alternatives that can provide information and correlations among industrial processes, wastewater characteristics and efficiencies of the wastewater treatment processes.

Some processes, e.g. industrial wastewater treatment, exhibit nonlinear behaviors which are difficult to describe using linear mathematical models. Therefore, the use of predictive models based on Artificial Neural Networks (ANNs), e.g. Multilayer Perceptron (MLP), to improve the operational control of wastewater treatment plants have been suggested in the literature [1–6]. Grieu et al. [7] presented a procedure based on a MLP network to predict influent and effluent ammonia and organic matter concentrations. The authors showed that neural modeling can be a useful tool to minimize operation costs and provide stability to the treatment process. More recent, Levenberg–Marquardt algorithm have been used for ANNs with a little or moderate training set (up to several hundred weights) because it requires a large storage memory for execution. It has been proven to be fast, convergent and robust [1].

An ANN normally relies on representative historical data of the process. Therefore, data preparation is an essential step for enhanced performance of these predictive models. This task requires a careful analysis of the data in order to define which variables best represent the system. Frequently, researchers encounter a large set of independent variables for possible inclusion in a multivariate analysis, but, in most cases, the inclusion of all variables in the model is unnecessary and can hinder the correct interpretation of the data.

Principal Component Analysis (PCA) is a multivariate statistical technique that reduces a complex system of correlations to a smaller number of dimensions. PCA can be used in three ways: to select principal components, to discard original variables and to exclude possible outliers. The main purpose of this technique is to reduce the dimensionality of a data-set consisting of a large number of interrelated variables while retaining as many of the variations present in the data-set as possible [8,9]. Considering a data matrix X with n rows (observations) and p columns (variables), Principal Components (PCs) are obtained by the diagonalization of the covariance matrix X^TX, where X^T is a transposed matrix of X. The elements of the eigenvectors, called *loadings* (weights) in PCA terminology, represent the cosine directors. In other words, they represent a contribution of each original axis in the new axis called principal components. The eigenvalues represent the amount of variance described by the original eigenvectors. PCs are uncorrelated and are ordered so that the first few retain most of the variations present in the original set. Therefore, the first PC (PC1) represents an axis in which the samples have maximum variance, i.e. an axis along which the samples have greater dispersion. PC2 represents a second axis with more variance, i.e. an axis with maximum variance not explained by PC1. This pattern is then repeated until the last PC.

K-Means clustering and PCA have been used to optimize the MLP learning phase [7]. Together these are able to minimize the overfitting phenomenon and help the identification of the strength of the variable correlations [7–9]. There are several methods for selecting the variables using PCA [10–13]. For example, Jolliffe [14] tested five methods: a multiple correlation method, two principal component methods and two clustering methods and proposed a method, designated B4, which discards the original variables based on the loading vectors of the first PCs. Several other studies in the literature have also successfully applied the PCA together with an ANN [15–18].

The purpose of this work is to predict organic matter removal from a pulp and paper mill wastewater treatment plant using a MLP Model and to evaluate the use of PCA as a preprocessing technique for selecting variables/PCs and identifying outliers.

2. Methods

2.1. Process description

The wastewater treatment plant consists of two steps: a preliminary physical–chemical treatment composed of two parallel mixing tanks and flocculation chambers followed by a biological treatment. The biological treatment is performed in an aerated lagoon (169,000 cubic meters) supplemented by a set of five stabilization ponds with a total volume of approximately 1.5 million cubic meters. The biological sludge is removed in settlement tanks and goes into a controlled landfill after the drying operation.

2.2. Data collection for prediction model

The original data base covered a period of 1,427 consecutive days, which represented daily recording for approximately 4 years. In order to minimize the loss of information due to exclusion of samples containing high incidence of missing values (>50%), the data-sets contained only the most frequently measured variables, i.e. flow rate (Q), influent organic matter (soluble COD_{in}), pH value, color, temperature, electrical conductivity (EC), wastewater flow coming from the pulp production (Q_{pulp}) and wastewater flow coming from the paper production (Q_{paper}). The biochemical oxygen demand (BOD) was not chosen as an input variable due to the significant amount of time for measurement (~5 d), which made it impractical to build the model. After eliminating samples that contain missing data or BOD and accounting for probable errors of measurement, the final data-set was reduced to 786 samples. Table 1 shows the basic statistical properties for the selected variables.

Five models were constructed to predict the content of soluble organic matter in the effluent of the aerated lagoon (COD_{out}). Model 1 (M1) was constructed by quantifying the organic matter present in the wastewater as a concentration of soluble COD (mg L^{-1}), while Models 2–5 were constructed using the organic load (COD_{load}) calculated by the multiplication of the COD concentration and flow rate. In Models 3–5, PCA was applied to reduce the dimensionality of the data-set in order to select PCs, discard original variables and exclude possible outliers.

2.3. ANN structure

The B4 method was used to discard the original variables based on the weight vectors of the first principal component. MLP was the ANN used for the prediction of the amount of organic matter effluent of the aerated lagoon (COD_{out}). The training algorithm was of the Levenberg–Marquardt type, which is an adaptation of the backpropagation algorithm. The neural network parameters can be changed to reach a suitable network architecture and to find a model with a more satisfactory result. The network parameters that were changed on the length of the training were learning rate, number of hidden layers and number of neurons per each hidden layer. The data-set was randomized and divided into the following three sets: training, validation and test sets. The transfer functions were log-sigmoid and linear for the intermediate and the output layer, respectively.

The linear activation function for the output neuron was appropriate for the continuous-variable targets. Sigmoidal activation functions for the input and hidden neurons were needed to introduce nonlinearity into the network. Without nonlinearity, hidden layers would not make the nets any more powerful than plain perceptrons (which do not have any hidden units and only contain input and output units). Sigmoidal activation functions are usually preferable to threshold activation functions [19].

2.4. Evaluation of the ANN model performance

The performance of each network model was evaluated by computing the mean square error (MSE), the linear correlation index (R^2) and the adjusted linear correlation index ($R^2_{adjusted}$). In contrast to R^2, the adjusted R^2 increases only if the additional model parameters significantly improve the regression results, which compensate for the increase in the regression degrees of freedom. Therefore, the adjusted R is the only statistical parameter able to perform reliable comparative analyses of the predictive performances of the ANN models. Minitab® and Matlab® were used for statistical analysis as well as for PCA and ANN modeling.

The examination of the adequacy of the model requires that the errors generated are normally dis-

Table 1
Basic statistical properties of the selected variables

Parameters	Mean	Standard deviation	Minimum	Maximum	Missing data (%)
Q (m^3 d^{-1})	67,364.0	11,588.0	4,474.0	97,850.0	0
Soluble COD_{in} (mg L^{-1})	562	104	136	925	6.2
pH	7.5	1.2	1.0	12.5	3.7
Color (units Pt-Co)	464.4	123.6	41	1,317	3.6
Temperature (°C)	45.5	3.1	28	50.5	32.6
EC (μS cm^{-1})	1,530.9	378.1	379	5,810	3.9
Q_{pulp} (ton d^{-1})	886.1	155.2	0	1,112.1	7.9
Q_{paper} (ton d^{-1})	1,042.7	94.2	382.4	1,304.8	6.5
Soluble COD_{out} (mg O_2 L^{-1})	315.5	2.0	105	865	5.8

tributed. We built normal residuals probability graphs and also standardized the residuals to estimate the normal distribution. During standardization, the residues were scaled to ensure that the values of the standard deviations were approximately equal to one. Therefore, large residuals (which may indicate possible outliers or unusual observations) would stand out from inspections of the residual plots. The residues were standardized using Eq. (1):

$$d_i = e_i/\sqrt{\sigma^2}, \quad i = 1, 2, \ldots, n \tag{1}$$

where e_i is the error calculated from the difference between the value predicted by the ANN and the desired output value, and σ is the standard deviation of the sample.

The criterion used to define the normal distribution of the data was that 95% of the standardized residuals were within the range (−2.2).

3. Results and discussion

Table 2 shows the variance and the weights of the principal components. According to the criteria of the B4 method, the first five components should be preserved in building the model because these PCs express 89.8% of the total preserved variance of the system. Using selected components, the main original variables were extracted from the absolute value of the loadings as proposed by Method B4. The most important variables (in the order of importance) were flow rate (Q), soluble COD_{load}, Q_{paper}, color, Q_{pulp} and temperature. Because soluble $COD_{load} = f(COD, Q)$ and the flow rate represent the same loading value, only soluble COD_{load} was kept as an input variable in the model. We note that the select variables from the B4 method were the same when using training, validation or test data-sets.

To understand how each variable impacts each component, graphic interpretation of the scores and loadings can be used. Fig. 1 shows the graph of the loadings of the components 1 and 2.

Analyzing the graph of the loadings, it can be seen that Q_{paper}, Q_{pulp} and temperature, as well as flow rate and COD_{load}, are correlated, as expected. The pH and EC may be considered as independent variables. The graph of the scores was used for identification of the outliers. We analyzed only the first PCs, as these individually retain the highest variability data. The questionable data are circled in Fig. 2.

As previously explained, the M5 model disregards these data to assess whether the PCA was useful for excluding outliers. The set of data used to generate the M5 model is composed of the original 8 variables and 719 values.

Table 3 shows the performance evaluation of each model. The best models were obtained when the ANNs were composed of only one hidden layer, the learning rate was equal to 0.05 and the division of the data to perform the training sets, validation and testing were equal to 70, 20 and 10%, respectively. By examining the values of R^2 and $R^2_{adjusted}$, only M1 showed poor performance. A better performance was found using models M2–M5 because they were built using the amount of influent organic matter expressed

Table 2
Variance and weights of the principal components

Principal components	Variance	Explained variance (%)	Accumulated variance (%)
PC_1	2.7	33.8	33.8
PC_2	1.7	20.6	54.4
PC_3	1.3	15.7	70.1
PC_4	0.8	10.6	80.7
PC_5	0.7	9.0	89.7
PC_6	0.4	5.5	95.2
PC_7	0.4	4.8	100
PC_8	0	0	100

Weights

Principal components	Q	COD_{load}	pH	Color	T	EC	Q_{pulp}	Q_{paper}
PC_1	*−0.49*	*−0.49*	0.27	−0.17	−0.28	0.38	−0.37	−0.24
PC_2	−0.42	−0.42	−0.35	−0.05	0.37	0.07	0.39	*0.48*
PC_3	0.06	0.06	0.53	*0.58*	0.32	0.49	0.16	0.10
PC_4	−0.11	−0.11	−0.26	0.70	−0.35	−0.22	*−0.39*	0.31
PC_5	−0.15	−0.15	−0.18	0.27	*0.58*	−0.29	−0.17	−0.64

Notes: Italic values represent the values of the loadings chosen as proposed by B4 Method.

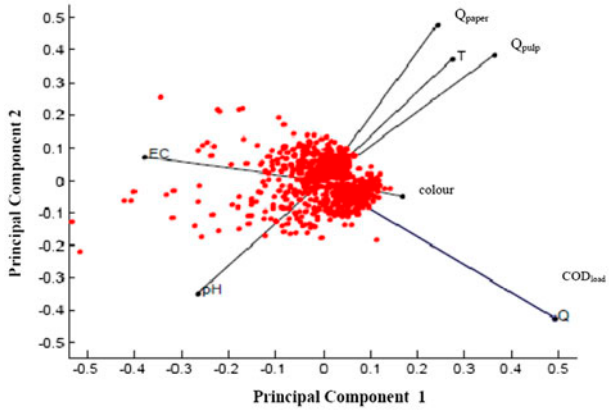

Fig. 1. Graph of the loadings of the two first PCs.

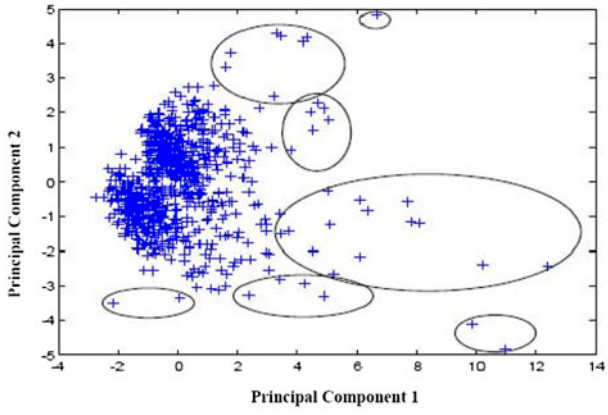

Fig. 2. Graph of the scores of the two first PCs.

in load terms (kg d^{-1}) instead of concentration terms. The Model M1 was built by considering the concentration of organic matter as COD and the flow rate. Therefore, the fluctuation of values from these parameters may have been the cause of M1's poor performance. Considering the models M2–M5, no significant differences were identified in the R^2 values. Previous attempts of modeling to predict organic matter utilized the COD concentration as an input variable, which resulted in R values ranging from 0.4 to 0.7 [1,4,5,7]. The lowest MSE values are obtained in M2 and M4 using the original variables as inputs. The number of iterations varied from model to model, but this was not a significant factor in affecting the current results because it does not take more than 10 s to run each model.

The adjusted R^2 values calculated using the data from models M2–M5 also showed no significant difference, which means that the network performance was unaffected by the reduction in adjustable parameters. Therefore, the complexity of the models can be reduced through appropriate data preparation. For example, M3 resulted in a faster learning (i.e. 18 interactions) when comparing the structure of the variability of the ANN models simply by switching from the original variables (M2) to the corresponding PCs (M3). This was expected because the M3 model was built with a smaller number of input data; however, the results showed no significant loss of information, which can be considered an advantage of using PCs. Furthermore, the same performance was repeated when we selected variables using the B4 method (M4).

Table 3
Evaluation of the models to predict the soluble COD_{out} of the aerated lagoon

	Models				
Comparative parameters	M1[a]	M2[b]	M3[b]	M4[b]	M5[b]
Inputs	8 original variables	8 original variables	5 PCs	5 original variable	8 original variables
Training data	706	706	706	706	706
Test data	80	80	80	80	80
Number of hidden neurons	1	1	1	1	1
Number of parameters	9	9	6	6	9
Transfer function	Sigmoid	Sigmoid	Sigmoid	Sigmoid	Sigmoid
Processing method	Back-propagation	Back-propagation	Back-propagation	Back-propagation	Back-propagation
Number of iterations	11	103	18	93	103
MSE test	2.3E-03	4.59E-08	1.9E-05	2.9E-08	4.0E-05
R^2 test	0.4508	0.9999	0.9953	0.9999	0.9999
Adjusted R^2	0.3753	0.9999	0.9796	0.9999	0.9999

[a]COD (mg O_2 L^{-1}).
[b]COD_{load} (kg d^{-1}).

Table 4
Results of the synaptic weights of the models

Inputs variable of the MLP									
Models	$COD_{soluble}$	Q	$COD_{soluble}$	pH	Color	T	EC	Q_{Pulp}	Q_{Paper}
M1	mg L^{-1}	−0.1514	−1.0645	0.1975	−0.0483	−0.0391	−0.3931	−0.0707	−0.0647
M2	kg d^{-1}	−0.1243	−0.3201	0.0001	0.0003	0.0001	−0.0003	0.0002	0.0001
M4	kg d^{-1}	–	−0.46	–	−0.003	0	–	−0.0002	−0.0001
M5	kg d^{-1}	−0.1243	−0.3201	0.0001	0.0003	0.0001	−0.0003	0.0002	0.0001
		PC_1	PC_2	PC_3	PC_4	PC_5			
M3	kg d^{-1}	−1.2451	−0.6821	0.1027	−0.1672	−0.1703	–	–	–

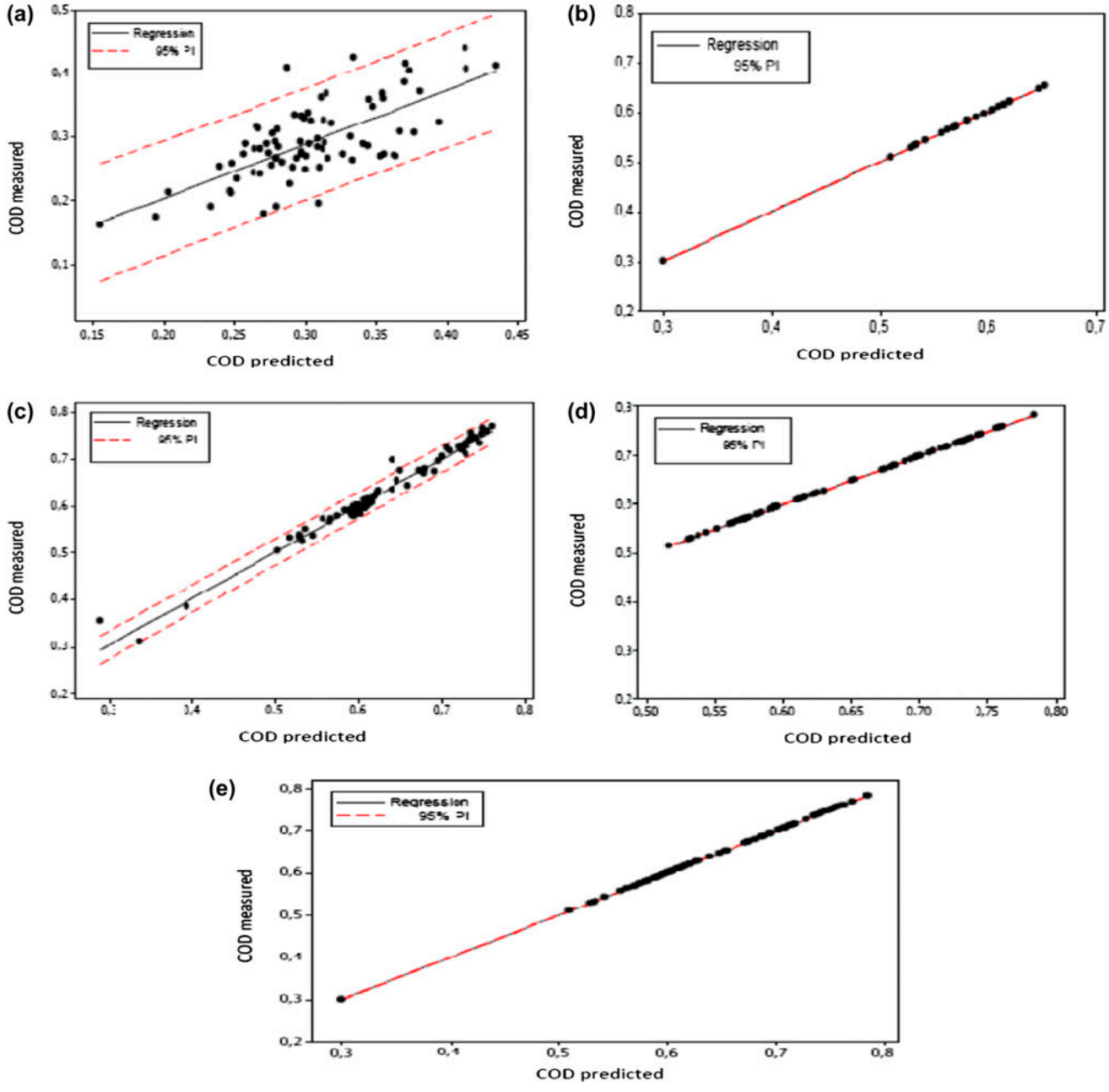

Fig. 3. Results of the Anderson–Darling normality test with a confidence level of 95% for (a) Model M1, (b) Model M2, (c) Model M3, (d) Model M4 and (e) Model M5 for dimensionless COD.

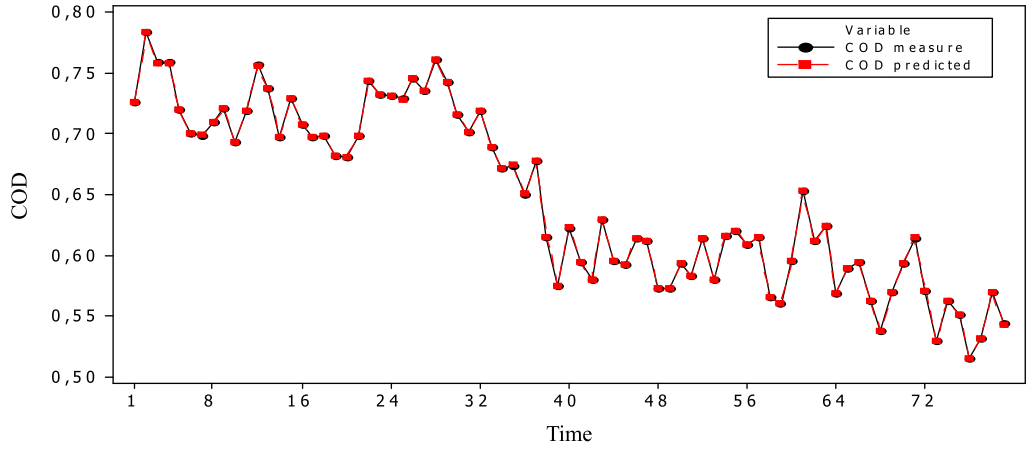

Fig. 4. Time series plot of the measured and predicted soluble COD_{out} of the M4 model.

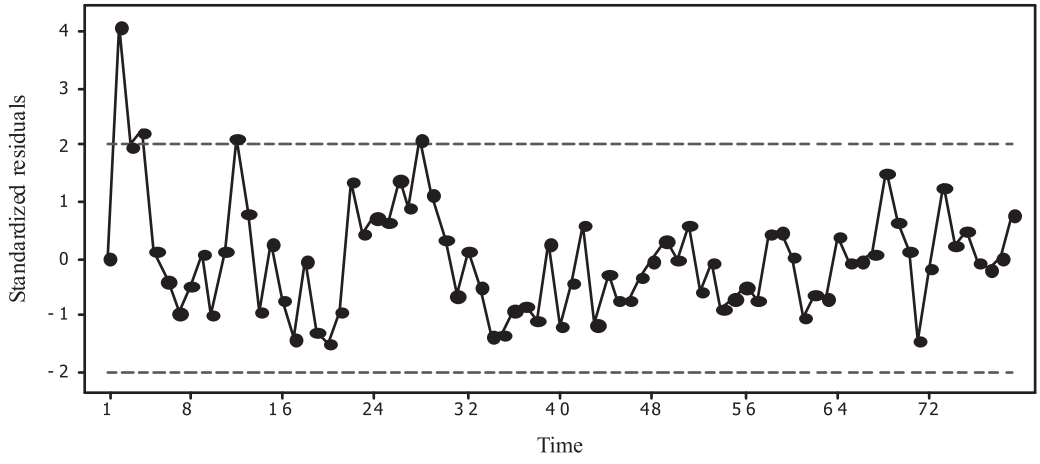

Fig. 5. Time series plot of the standardized residuals of the M4 model.

Therefore, the signal-to-noise ratio does not affect the PCA application in this case.

The error rate is generally more significant because it is a supervised neural network type. It should be noted that the model M3 was built using only five PCs as predictor variables in the ANN but required information from eight original variables. The results obtained using the PCA to exclude possible outliers in M5 were similar to that of M3, which means that the exclusion of outliers is unnecessary in this case. However, this result cannot be generalized. In fact, the presence of outliers can provide incorrect or misleading results, mainly during the construction of the empirical models. The model M4 was built by discarding pH and EC variables; hence, it was the most synthetic and the simplest model obtained. Thus, it can be concluded that the two variables discarded do not add information nor influence the performance of the prediction model.

Table 4 shows the synaptic weights related to each input variable of the models. The results indicate that the synaptic weights of the variables, influent flow rate and COD, were significant in building the models M2, M4 and M5. In contrast, in models M1 and M3, the weights of all variables are in the same order of magnitude. This result indicates that some information, such as the COD concentration and the use of PCs as input variables of the MLP, were satisfactory for predicting the organic matter effluent of the aerated lagoon. However, for a better understanding of synaptic weights, it is necessary to perform a sensitivity analysis, which was not conducted in this work.

Eq. (2) below is obtained from the model M4 and predicts the removal of the organic matter from pulp

and paper mill wastewater. Since it is a empirical model, it cannot be used to represent other systems.

$$\mathrm{COD_{out}} = \frac{8.67}{1 + e^{-0.46\,\mathrm{COD_{in}} - 0.003\,\mathrm{Color} - 0.0002\,Q_{\mathrm{paper}} - 0.001\,Q_{\mathrm{pulp}} + 0.0017}} - 4.3 \quad (2)$$

This model eliminates the temperature (T) variable and uses four dependent variables to obtain the outgoing organic load of the aerated lagoon. Fig. 3 shows the graphs of the effluent soluble COD determined in laboratory tests vs. the predicted values using the Anderson–Darling normality test with a confidence level of 95%. It can be seen that the models M4 and M2 are able to predict the behavior of the bioprocess perfectly. Furthermore, as explained previously, the model M4 was built by discarding two variables. Hence, it can be concluded that M4 is the best model obtained in the current study.

Fig. 4 shows the comparison between predicted and measured values of the M4 test set. It is observed that this model perfectly reproduces the overall variation observed in the biological treatment. Fig. 5 shows the time series plots of the standardized residuals of the M4 model. It is noted here that 95% of the residuals fall in the range (−2, +2), confirming the normality of residuals and the adequacy of the model. Standardization was carried out using the standard deviation of the validation data-set.

The methodology presented in this work can be used for other data and processes. The basic regression procedure using neural networks follows a well-defined routine. The greatest variation occurs only with the pre-treatment data.

4. Conclusions

This research investigated the use of PCA as a data-preprocessing technique to build an ANN model in order to predict organic matter removal from pulp and paper mill wastewater. It was concluded that PCA, which was applied to select input variables, can be useful in neural network learning processes. The use of this technique allowed for the reduction of the number of parameters to be adjusted without changing the performance of the model. Furthermore, utilizing the PCA to discard original variables made it possible to improve neural network performance without any loss of information. The use of an ANN model may reduce costs by discarding unnecessary laboratory measurements. However, in this particular case, the PCA technique proved to be unnecessary for outlier exclusion.

It is important to note that choosing the best ANN model should not be done arbitrarily and carelessly. It is necessary to use various statistical parameters to compare models of different sizes and structures in order to determine the best model. In order for the preprocessing data to be meaningful, it is strongly recommended that a professional who has expertise in the process should be available for examination.

Acknowledgments

The authors would like to thank the Foundation for Research Support of the Bahia State for financial and material support.

Nomenclature

R^2_{adjusted}	—	adjusted linear correlation index
Q	—	affluent flow rate
ANN	—	artificial neural network
BOD	—	biochemical oxygen demand
COD	—	chemical oxygen demand
Colour	—	color
EC	—	electric conductivity
MSE	—	MEAN square error
Q_{paper}	—	paper production flow rate
pH	—	pH
PC	—	principal component
PCA	—	principal component analysis
R^2	—	linear correlation index
MLP	—	multilayer perceptron
Q_{pulp}	—	pulp production flow rate
σ	—	standard deviation
T	—	temperature

References

[1] M.M. Hamed, M.G. Khalafallah, E.A. Hassanien, Prediction of wastewater treatment plant performance using artificial neural networks, Environ. Modell. Software 19 (2004) 919–928.
[2] F.S. Mjalli, S. Al-Asheh, H.E. Alfadala, Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance, J. Environ. Manage. 83 (2007) 329–338.
[3] H. Chen, S. Ning, R. Yu, M. Hung, Optimizing the monitoring strategy of wastewater treatment plants by multiobjective neural networks approach, Environ. Monit. Assess. 125 (2007) 325–332.
[4] C.S. Akratos, J.N.E. Papaspyros, V.A. Tsihrintzis, An artificial neural network model and design equations for BOD and COD removal prediction in horizontal subsurface flow constructed wetlands, Chem. Eng. J. 143 (2008) 96–110.

[5] D.B. May, M. Sivakumar, Prediction of urban stormwater quality using artificial neural networks, Environ. Modell. Software 24 (2009) 296–302.
[6] R. Mohammadi, H. Eskandarloo, M. Mohammadi, Application of artificial neural network (ANN) for modeling of dyes decolorization by Sn/Zn-TiO_2 nanoparticles, Desalin. Water Treat. 55 (2015) 1922–1933.
[7] S. Grieu, A. Traore, M. Polit, J. Colprim, Prediction of parameters characterizing the state of a pollution removal biologic process, Eng. Appl. Artif. Intell. 18 (2005) 559–573.
[8] T. Kourti, P. Nomikos, J.F. MacGregor, Process analysis, monitoring and diagnosis using multivariate projection methods, Chemom. Intell. Lab. Syst. 28 (1995) 3–21.
[9] A.P.D. Silva, Efficient variable screening for multivariate analysis, J. Multivariate Anal. 76 (2001) 35–62.
[10] J.A. Cumming, D.A. Wooff, Dimension reduction via principal variables, Comput. Stat. Data Anal. 52 (2007) 550–565.
[11] J. Cadima, J.O. Cerdeira, M. Minhoto, Computational aspects of algorithms for variable selection in the context of principal components, Comput. Stat. Data Anal. 47 (2008) 225–236.
[12] K. Fueda, M. Iizuka, Y. Mori, Variable selection in multivariate methods using global score estimation, Comput. Stat. 24 (2009) 127–144.
[13] M. Matteau, A.A. Assani, M. Mesfioui, Application of multivariate statistical analysis methods to the dam hydrologic impact studies, J. Hydrol. 371 (2009) 120–128.
[14] I.T. Jolliffe, Discarding variables in a principal component analysis. I: Artificial data, Appl. Stat-J. Roy. St. C. 2 (1972) 160–173.
[15] A. Abbaspour, L. Baramakeh, Application of principle component analysis artificial neural network for simultaneous determination of zirconium and hafnium in real samples, Spectrochim. Acta Part A: Molecular and Biomolecular Spectroscopy 64 (2006) 477–482.
[16] F.H. Bellamine, A. Elkamel, Model order reduction using neural network principal component analysis and generalized dimensional analysis, Eng. Comput. 25 (2008) 443–463.
[17] C.L. Wu, K.W. Chau, C. Fan, Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques, J. Hydrol. 389 (2010) 146–167.
[18] M. Palit, B. Tudu, N. Bhattacharyya, A. Dutta, P.K. Dutta, A. Jana, R. Bandyopadhyay, A. Chatterjee, Comparison of multivariate preprocessing techniques as applied to electronic tongue based pattern classification for black tea, Anal. Chim. Acta 675 (2010) 8–15.
[19] R. Saraswathi, M.K. Saseetharan, S. Suja, ANN–based predictive model for performance evaluation of paper and pulp effluent treatment plant, Int. J. Comput. Appl. Technol. 45 (2012) 280–289.