ELSEVIER

Chemical Engineering Journal 143 (2008) 96–110

Chemical Engineering Journal

www.elsevier.com/locate/cej

An artificial neural network model and design equations for BOD and COD removal prediction in horizontal subsurface flow constructed wetlands

Christos S. Akratos, John N.E. Papaspyros, Vassilios A. Tsihrintzis *

Laboratory of Ecological Engineering and Technology, Department of Environmental Engineering, School of Engineering, Democritus University of Thrace, 67100 Xanthi, Greece

Received 18 December 2006; received in revised form 15 December 2007; accepted 20 December 2007

Abstract

A model is presented, which can be used in the design of horizontal subsurface flow (HSF) constructed wetlands. This model was developed based on experimental data from five pilot-scale CW units, used in conjunction with artificial neural networks. The CWs were operated for a two-year period under four different hydraulic residence times (HRT). For the proper selection of the parameters entering the neural network, a principal component analysis (PCA) was performed first. From the PCA and model results, it occurs that the main parameters affecting BOD removal are porous media porosity, wastewater temperature and hydraulic residence time, and a set of other parameters which include the meteorological ones. Two artificial neural networks (ANNs) were examined: the first included only the three main parameters selected from the PCA, and the second included, in addition, the meteorological parameters. The first ANN predicted BOD removal rather satisfactorily and the second one examined resulted in even better predictions. From the predictions of the ANNs, a hyperbolic design equation, which combines zero and first order kinetics, was produced to predict BOD removal. The results of the ANNs and of the model design equation were compared to available data from the literature, and showed a rather satisfactory correlation. COD removal was found to be strongly correlated to BOD removal. An equation for COD removal prediction was also produced.

Keywords: Constructed wetlands; Horizontal subsurface flow; Principal component analysis; Artificial neural networks; BOD; COD; Removal model

1. Introduction and background

The use of constructed wetlands (CWs) in the treatment of a variety of wastewaters has grown rapidly since the mid 1980s. Several advantages of CWs make them promising as a low cost alternative for wastewater treatment. In conjunction with EU directive 1991/271/EEC, which makes municipal wastewater treatment imperative even for small towns, CWs are classified as a key technology. The main characteristics, which affect the removal efficiency of CWs, are the hydraulic residence time and temperature [1].

The effectiveness of CWs in removing organic matter, nitrogen and phosphorus has been proved by several studies [e.g., 2–5]. Organic matter is removed in a horizontal subsurface flow (HSF) constructed wetland by bacteria attached to the porous media and plant roots [6–9]. These bacteria remove organic compounds with several mechanisms (i.e., respiration, fermentation and methanogenesis). Plant roots provide the necessary surfaces for bacteria to grow and supply oxygen to them [10].

Until now, the majority of the models on constructed wetlands are focused on input–output data and the production of either linear regression equations or first order decay models [11]. Although, regression equations are very useful, they simplify a complex system, such as constructed wetlands, into only two or three parameters, ignoring important factors such as climate, porous media material and plant type [11]. The most common type of models used in constructed wetland design are the first order equations, which predict an exponential decay between inlet and outlet concentrations under constant conditions of influent [11]. The main drawbacks of the first order equations are that they assume an ideal and constant flow [11], and they do not take into consideration porous media, plant and climate. Table 1 presents the main regression and first order equations for predicting BOD removal in HSF CWs, as pre-

* Corresponding author. Tel.: +30 25410 79393; fax: +30 25410 78113.
E-mail addresses: tsihrin@otenet.gr, tsihrin@env.duth.gr (V.A. Tsihrintzis).

1385-8947/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cej.2007.12.029

Table 1
Main literature regression and first order equations for BOD/COD removal in CWs [11]

Regression equations

Reference	Equation	Influent range	Effluent range	q Range
[31]	$C_{out} = (0.11C_{in}) + 1.87$	$1 < C_{in} < 330$	$1 < C_{out} < 50$	$0.8 < q < 22$
[32]	$C_{out} = 0.33C_{in} + 14$	$1 < C_{in} < 57$	$1 < C_{out} < 36$	$1.9 < q < 11.4$
[33]	$C_{out} = 502.20e^{-0.111T}$	$10 < T < 30$	–	–
[34]	$C_{out} = (0.099C_{in}) + 3.24$	$5.8 < C_{in} < 328$	$5.8 < C_{out} < 51$	$0.6 < q < 14.2$
[26]	$L_{removed} = (0.653L_{in}) + 0.292$	$4 < L_{in} < 145$	$4 < L_{removed} < 88$	–
[34]	$L_{out} = (0.145L_{in}) - 0.06$	$6 < L_{in} < 76$	$0.3 < L_{out} < 11$	–
[34]	$L_{out} = (0.13L_{in}) + 0.27$	$2.6 < L_{in} < 99.6$	$0.32 < L_{out} < 21.7$	$0.6 < q < 14.2$

First order equations

Reference	Eq. (1)			Eq. (2)		
	k_A (m/d)	$k_{A,20}$ (m/d)	θ	k_v (d^{-1})	$k_{v,20}$(d^{-1})	θ
BOD_5						
[3]					1.104	1.06
[21]				0.17		
[21]					0.22	1.06
[35]				1.84		
[35]				1.35		
[35]				0.86		
[4]	0.085–1.000			0.3–6.11		
[36]		0.49	1.0			
[37]		0.86	1.1			
[38]	0.118 ± 0.022					
[39]	0.083					
[40]	0.067 ± 0.1					
[31]	0.16					
[31]	0.068					
[41]	0.133					
[41]	0.07–0.31					
[40]	0.06					
[40]	0.31					
[41]	0.17					
[42]				0.86		
COD						
[5]					0.031	1.06

Note: C_{in} = influent concentration in mg/L; C_{out} = effluent concentration in mg/L; L_{in} = influent surface loading rate in kg/m^2/d; L_{out} = effluent surface loading rate in kg/m^2/d; $L_{removed}$ = removed surface loading rate in kg/m^2/d; T = temperature in °C; q = hydraulic surface loading rate in m^3/m^2/d.

sented by Rousseau et al. [11]. A common form of the first order equations is presented by [4]:

$$\frac{C_{out}}{C_{in}} = e^{(-k_A/q)} \quad (1)$$

where q is the hydraulic loading rate in m/d and k_A the decomposition constant in m/d.

Another form of the first order equation, which uses the hydraulic residence time (HRT) t in days, is presented by Eq. (2) [3]:

$$\frac{C_{out}}{C_{in}} = e^{-k_v t} \quad (2)$$

The temperature effect is expressed by the constant k_T (k_A or k_v), which is determined by the use of an Arrhenious equation, as follows [4]:

$$k_T = k_{20}\theta^{(T-20)} \quad (3)$$

where k_{20} ($k_{A,20}$ or $k_{v,20}$) is the value of k_T at 20 °C. The k_{20} and θ constants usually result from statistical analysis of the data used in the production of the model.

A critical step in model production is the selection of the variables, which can describe removal efficiency satisfactorily. Principal component analysis (PCA) is a multivariate linear technique useful in data reduction, which enables highly correlated variables to be reduced to a small number of orthogonal parameters [12]. PCA can be applied to both continuous and categorical variables ("leveled" variables, like month etc.) [13]. In a geometric metaphor of PCA, a cloud of data points (in this study, BOD removal values) in a space of N input variables is reoriented so that the first axis (i.e., the first "principal component") is along the longest dimension of the cloud (thereby explaining most of the variance of the data cloud), the second axis is along the second longest dimension that is perpendicular to the first axis, and so on. Then, the original N input variables are examined with respect to their projection on the new axes. Input

variables with close projections on the new axis system are considered to be correlated (in which case only one of them could be used to model the data cloud), and variables with greatest projections along one of the new axes are considered to be the most strongly related to the corresponding principal component (only the variable with the greatest projection on a principal component axis is used in the subsequent analysis). A typical criterion for the number of principal components of a dataset is explanation of 75–95% of total variance of the dataset by the principal components [14].

ANNs are a technique inspired by biological neuron processing. They have a wide application field on several sciences for time series forecasting, pattern recognition and process control [15]. Their main advantage over traditional methods is that they do not require the complex nature of the underlying process [15]. The principal drawback of ANNs is that they are typically used as a "black-box" approach, hiding the physics of the modeled process; in the present work, however, a model, inspired from the ANN response curves to the input parameters, is proposed as an alternative to the ANNs. This model successfully describes their complex dynamics. There are many types of networks for an ANN application and the selection of the proper type depends on the nature of the problem and data availability. The multi-layer perception (MLP) is perhaps the most popular network used in hydrological modeling [16,17]. In MLP, the artificial neurons, or processing units, are arranged in a layered configuration containing an input layer, a processing ("hidden") layer (in complex topologies two hidden layers are used) and an output layer.

A simple MLP was used in this work; it is a network with three input variables, a hidden layer with three processing neurons, and a single-unit output (i.e., there is only one output variable in this network, BOD removal). For a simple regression analysis the units in the input layer introduce normalized or filtered values of each input variable into the network, then these values are transferred to all units of the hidden layer multiplied by a "weight" factor that is, in general, different for every connection, and its magnitude characterizes the importance of some connection. The units of the hidden layer apply non-linear operators ("activation functions") on their input, which offer to the ANN its non-linear modeling capability, and the results of the non-linear operations are combined (again with proper weights for every connection) to the output layer. This is a feed-forward scheme, but "recurrent" interconnections (where some unit in a layer returns its output to an earlier layer) may also be used. Along with the network topology (i.e., number of layers and units in every layer, activation functions of the units, interconnection scheme), the weights of the connections are the design parameters of an ANN model. The weights are updated during the "training" of the model: in the so-called "supervised learning" training, the network output values are compared to desired output values (like experimental results, here BOD removal), and the weights are updated in a loop, until the network outputs are close enough to the desired output values. A classic algorithm for the update of the weights is the "back-propagation" algorithm.

The aim of the present study is to examine whether artificial neural networks (ANNs) could be used in predicting BOD and COD removals in horizontal subsurface flow constructed wetlands, and if so, to suggest an appropriate topology (i.e., input variables, number of ANN neurons, training algorithm, etc.) for a successful ANN. The data used to train and test the ANN come from five pilot-scale horizontal subsurface flow constructed wetlands receiving synthetic wastewater, which were operated for two years in controlled experiments under Mediterranean climate conditions. To the authors' knowledge, there are no publications studying the applicability of ANNs in predicting the performance of constructed wetlands. Another novelty of the present work is that the ANN results are used as an aid in the development of a design equation for BOD and COD removals. The design equation proposed here is a tool complementary to ANN analysis, and it will be shown to be a simplifying form of the first-order decay law.

2. Materials and methods

2.1. Pilot-scale unit description

Five similar pilot-scale horizontal subsurface flow constructed wetlands have been constructed and are in operation. A detailed description of the experimental facility has been presented by Akratos and Tsihrintzis [18]. Briefly, they are rectangular tanks made of steel, with dimensions 3 m long, 0.75 m wide and 1 m deep. Three different porous media were used, i.e., medium gravel (MG, D_{50} = 15.0 mm, range 4–25 mm), fine gravel (FG, D_{50} = 6 mm, range 0.25–16.0 mm) and cobbles (CO, D_{50} = 90 mm, range 30–180 mm). The porous media were of different origin. Medium gravel, which was obtained from a quarry, is a carbonate rock (main elements: Si 3.39%; Al 0.90%; Fe 0.82%; Ca 27.20%; Mg 4.53%; P 0.03%). Fine gravel and cobbles were both obtained from a river bed in the area, and are igneous rocks (Si 28.50%; Al 7.95%; Fe 4.22%; Ca 3.62%; Mg 1.76%; P 0.11%). The rest of the rock mass consists of various trace elements in very small quantities. Two plants were used, namely common reed (R, *Phragmites australis*) and cattails (C, *Typha latifolia*). The plants were collected from watercourses in the vicinity of the laboratory. Medium gravel from the quarry was used in three wetland units, one planted with common reed (MG-R) and one with cattail (MG-C); the third one was kept unplanted (MG-Z). The other two units contained fine gravel (FG-R) and cobbles (CO-R) from the river bed, both planted with common reed. Synthetic wastewater was designed and used to simulate to the best the characteristics of domestic wastewater [18]. Inflowing concentrations of BOD and COD were approximately 300 mg/L and 600 mg/L, respectively. The synthetic wastewater contained organic substances and a source of nitrogen, phosphorus and trace elements. The range of the flow was from 16 to 55 L/d in each unit, with four residence times, i.e., 6, 8, 14 and 20 days. The range of surface organic loading was: for BOD 2.6–8.8 g/d/m^2, for COD 4.2–14.3 g/d/m^2, for *ortho*-phosphate (OP) 0.06–0.18 mg/d/m^2 and for TKN 0.46–1.60 mg/d/m^2. The organic substances used were pepton (200 mg/L) which is the protein source, cane sugar

Table 2
Statistics of overall influent and effluent concentrations and removal efficiencies in each unit [18]

		Influent concentration (mg/L)	Effluent concentration (mg/L)					Removal efficiency (%)				
			MG-C	MG-R	MG-Z	FG-R	CO-R	MG-C	MG-R	MG-Z	FG-R	CO-R
BOD (mg/L)	Mean	361.1	41.2	53.9	50.1	38.8	45.2	88.3	84.6	85.7	89.0	87.0
	S.D.	47.4	33.1	37.6	34.3	26.9	36.3	10.3	11.6	10.5	8.5	11.6
	Min	282.0	0.6	7.0	4.0	5.1	4.0	49.1	46.1	53.3	47.4	31.7
	Max	507.0	170.0	190.0	140.0	169.0	200.0	99.8	98.2	98.9	98.4	98.9
COD (mg/L)	Mean	583.6	62.5	87.9	74.9	61.0	76.6	89.3	84.9	87.2	89.5	86.7
	S.D.	47.3	40.2	54.5	45.3	34.7	44.5	7.0	9.3	7.7	6.0	8.0
	Min	500.0	0.0	0.0	0.0	0.0	4.8	66.1	55.9	61.7	67.5	59.9
	Max	700.0	186.4	259.2	198.8	178.8	220.8	100.0	100.0	100.0	100.0	99.3
TKN (mg/L)	Mean	64.0	21.1	28.9	42.6	11.2	16.7	66.8	54.8	34.2	82.5	74.1
	S.D.	6.1	14.0	16.9	10.6	9.5	13.9	22.5	25.8	14.9	14.4	21.0
	Min	50.4	0.3	3.9	21.6	0.8	4.2	16.7	−2.3	0.8	44.6	7.3
	Max	77.0	53.2	75.9	65.5	37.8	64.9	99.4	92.8	70.5	98.8	94.2
$N\text{-}NH_3$ (mg/L)	Mean	38.4	17.1	24.3	38.3	7.5	10.8	53.6	36.2	−0.2	79.1	70.6
	S.D.	3.6	13.2	15.6	9.2	9.0	11.7	35.1	40.2	22.4	25.1	31.5
	Min	30.2	0.0	0.0	14.6	0.0	0.0	−20.3	−48.1	−63.8	0.0	−3.7
	Max	46.2	51.8	56.0	56.8	34.7	43.4	100.0	100.0	60.3	100.0	100.0
$N\text{-}(NO_3^- + NO_2^-)$ (μg N/L)	Mean	279.2	248.4	60.7	28.4	12.0	22.9					
	S.D.	379.8	495.8	162.4	119.4	29.7	56.7					
	Min	0.0	0.0	0.0	0.0	0.0	0.0					
	Max	1510.8	2068.0	748.7	752.0	138.9	280.8					
$P\text{–}PO_4^{3-}$ (mg/L)	Mean	8.2	2.7	5.8	4.5	0.9	3.5	66.9	28.2	43.9	88.6	57.3
	S.D.	1.0	2.7	2.5	1.5	1.4	2.5	33.8	30.7	19.1	17.7	30.4
	Min	6.0	0.0	1.2	1.8	0.0	0.0	−16.7	−54.1	1.2	14.1	−8.3
	Max	10.7	9.8	11.5	8.4	7.3	9.1	100.0	86.0	78.8	100.0	100.0
TP (mg/L)	Mean	9.1	3.8	7.2	5.2	1.6	5.1	58.4	21.1	42.6	81.5	44.5
	S.D.	1.0	3.4	2.8	1.5	2.1	2.8	36.3	29.2	17.0	24.0	29.1
	Min	7.7	0.0	2.0	2.5	0.0	0.0	−48.4	−51.6	2.4	6.1	−12.9
	Max	13.0	13.8	14.1	8.1	8.5	10.5	100.0	81.3	73.1	100.0	100.0

Number of Samples: 67 (54 for TP).

(200 mg/L) which is the source of saccharace, and acetic acid (50 mg/L) which is the source of organic acids. The source of phosphorus was hydrogen potassium phosphate (K_2HPO_4), with a typical inlet concentration of 10 mg/L PO_4^{3-}–P. The source of nitrogen was urea with a typical inlet concentration 60 mg/L NH_4^+–N. For the trace elements a fertilizer was chosen for use. The typical inlet concentration of trace elements were: Mg^{+2} 20 mg/L, Ca^{+2} 20 mg/L, Fe^{+2} <1 mg/L and W^{+2} <0.1 mg/L. The synthetic wastewater was produced three times a day, every eight hours, and was introduced immediately after mixing into the five tanks.

The data used in this paper come from the two-year operation of the pilot-scale units described above. Pollutant concentrations and removal statistics are presented in Table 2 [18].

2.2. PCA and neural networks software

All data analysis in this work was performed in StatSoft Statistica version 7 [13]. For PCA and ANN analysis, a linear transformation ($x' = (x–\mu)/\sigma$) was applied to every variable in the dataset, where x' is the transformed data value and x is the original data value, while μ and σ are the mean value and the standard deviation of the original data values, respectively. This transformation is called "soft scaling", and delivers transformed variables with zero mean and unit standard deviation. Such a transformation is critical for variables with values spanning over different ranges, to ensure that all variables have common importance in PCA and ANN analysis, and that the activation functions of the ANN units can operate on the input data, since several of these functions have an input range between −1 and 1, or between 0 and 1.

A modern algorithm, NIPALS, was used for PCA; this algorithm is considered numerically more accurate than other methods [13]. NIPALS is an iterative algorithm, finding one principal component at a time, by iterative regression of the projections of the original data points on the principal component axis, until the axis with the best "loading factors" is determined. The "loading factor" of some original variable with respect to some principal component axis is the cosine of the angle between the original variable axis and the principal component axis in the N-space of the original input variables; accordingly, the loading factor measures the contribution of a variable to a principal component. The "best" loading factors correspond to the line that explains most of the original data cloud variability.

For data analyses performed on samples, a "validation" procedure is required to ensure that the results of the analyses are of general validity, i.e., they do not hold just on the samples studied but on the entire population as well. A classic approach for validation is the division of the data to a "training" and a "validation" subset; analysis is performed on the training subset, and its results are checked in the validation subset. A different technique, namely cross-validation, is used in this work, as it is considered more trust-worthy than the classic approach, and it does not require an unbiased division of the original data to ensure that every subset of the data is representative of the original dataset. For PCA, the v-fold cross-validation was used: in this algorithm, data are divided into v segments, with $v - 1$ segments used to determine the principal axes and the last segment used for validation; the process is repeated for every possible permutation of the training and testing segments. PCA was performed with a 50-iteration maximum and a convergence criterion of 0.0001 for the iterative NIPALS algorithm, while $v = 7$ was used for the v-fold cross-validation.

For the ANN analysis, the choice today is usually between multiple layer perceptrons (MLPs) and Radial Basis Function (RBF) networks, which have a single hidden layer of radial units, with each unit modeling a Gaussian response surface. RBF networks have some advantages over MLPs, including faster training and less danger to hit local minima during training (recall that ANN training is an iterative process; thereby, it is possible for this process to stuck to some combination of the ANN weights that will give a local optimum for the ANN performance, instead of the global one). On the other hand, more units are typically required for the hidden layer of RBF networks as compared to MLPs [13]: this large number of units offers great flexibility to RBF networks, but exposes the analysis to the danger of "over-fitting". Over-fitting is one of the plagues of ANN analysis, and it could be described as a situation where too many units are used in an ANN, resulting to excellent performance of the ANN on the training data and striking failure on any data not used in the training process. In this work, feed-forward MLPs are used, with a single hidden layer, and a single output unit (BOD). To handle over-fitting, input data are split to three subsets in a 2:1:1 ratio, with the larger subset used for training, and the remaining two subsets used for "testing" and validation. During the "testing" phase, which runs parallel to training, the ANN delivered by training is applied to the testing subset, which does not participate in the training process: if the network "error" (i.e., the difference between the desired output and the network output) keeps dropping in the training subset but not in the testing subset, over-fitting is diagnosed and the training process is halted.

Training of the ANNs in this work was performed in two phases. During the first phase, the traditional back-propagation algorithm was run for 100 "epochs" (an epoch is one run of the network for the data points of the training subset, followed by one run for the data points in the testing subset to diagnose if over-fitting was present), with a "learning rate" of 0.01 (the learning rate is a control parameter for the step size of the adjustment of the network weights). Using another geometric metaphor for this first phase of training, in each epoch the back-propagation algorithm calculated the direction of the steepest descent on the error surface of the network, and jumped down the surface a distance proportional to the learning rate and the slope [13]. The first phase of training delivered an initial solution for the second phase, which was a more sophisticated conjugate gradient descent algorithm that ran for 500 epochs. This training method is suggested to be very efficient [13], and it was found to be so in the present work.

Concerning the activation function of the ANN units, a sigmoid function is typically used in MLP networks. The reason for this choice is that the sigmoid function is extremely flexi-

ble, combining nearly linear behavior, curvilinear behavior and nearly constant behavior, depending on the input value. In this work, a hyperbolic tangent function (tan h) is used; this is a variant of the sigmoid function, which takes values between −1 and 1 instead of between 0 and 1 of the sigmoid, and it is evaluated as

$$y(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (4)$$

Furthermore, a linear combination of the hidden layer values is used in the output unit; the alternative of using a sigmoid activation function for this unit was not found to be necessary.

A Monte-Carlo technique was used for cross-validation of the ANNs (that is, to ensure that the networks delivered will correspond to global, instead of local, minima of their error functions). The results of the neural network analyses of the present work are ensembles of several ANNs, with each member of an ensemble corresponding to a different and random split of the input data to 2:1:1 subsets, as mentioned above. The outputs of the member networks of some ensemble are averaged to provide a final, composite network prediction. It can be shown that the various performance indices of such an ensemble are conservative, i.e., any performance index of such an ensemble will be a "worst case" value [13]. A number of 50 ANNs per ensemble was found to be enough to guarantee cross-validation, and this is the number of ensemble members used in this work. It is stressed here that although training of the ensembles will require much more time compared to a single ANN, their practical response will still be very fast.

Finally, the BOD and COD removals in this work was tackled as a simple regression problem for ANN analysis, instead of a time series problem (which would result to more complex networks, requiring additional input parameters to include history information during both training and application of the models). This decision was supported by the fast response of the constructed wetlands studied here to changes of the primary input variables that affect BOD removal (that is, temperature and hydraulic residence time).

3. ANN for BOD removal

3.1. BOD removal data

BOD (5-day value) removal is defined as:

$$R_{BOD} = \frac{C_{in} - C_{out}}{C_{in}} \quad (5)$$

A total of 812 data points for BOD removal were collected and used in the analysis. BOD removal data were collected for hydraulic residence time (HRT) values of 6, 8, 14, and 20 days at the outflow of the units; in addition to these, the data collected at the 1/3 and 2/3 of the length of the units were also used, resulting in an HRT range of 2–days, thus covering most practical applications.

BOD removal efficiency over wastewater temperature is shown in Fig. 1a for three HRT values, i.e., 6, 14, and 20 days. Data from all units with the same HRT value were used. Representative third degree polynomials are also shown in Fig. 1a, one polynomial for every HRT value. The data look coherent, that is, there is a certain functional dependence on wastewater temperature for every HRT, with reasonably low scatter.

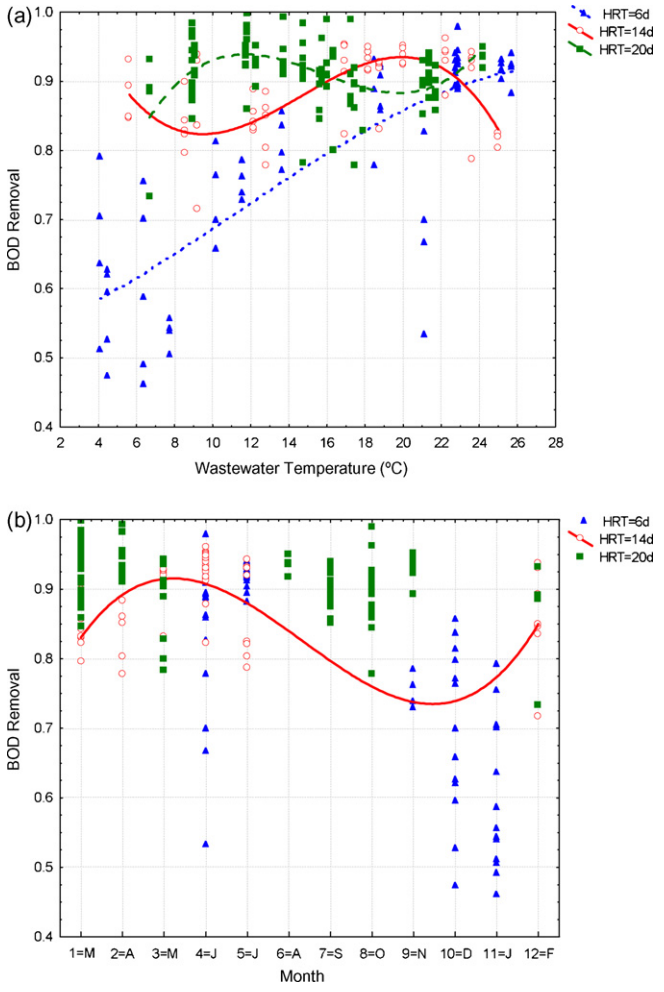

Fig. 1. (a) BOD removal efficiency chart over water temperature; (b) BOD removal efficiency chart over month.

A phase shift is apparent between the fits for HRT = 14 days (operated from February 2004 to July 2004) and HRT = 20 days (operated from February 2005 to May 2005 and August 2005 to November 2005), which is explained in Fig. 1b, where BOD removal performance is depicted over month instead of temperature. This implies that wetland performance should not be tested for temperature effect by graphing over temperature alone, but seasonal effects should also be considered. Month numbering in Fig. 1b starts from the first month of Spring (which is March in the present study; this is the month when stems and leaves of the plants emerge), as this would provide a more universal reference basis if the present data are compared to others of different seasonality. This trick was also found to offer better ANN modeling. There is a third degree polynomial shown in Fig. 1b (for HRT = 14 days), which is considered representative of the seasonal effect; accordingly, a sinusoidal function can be used to describe seasonal effects, as also suggested by Kadlec et al. [19]. It should be noticed that BOD removal peaks at summer start,

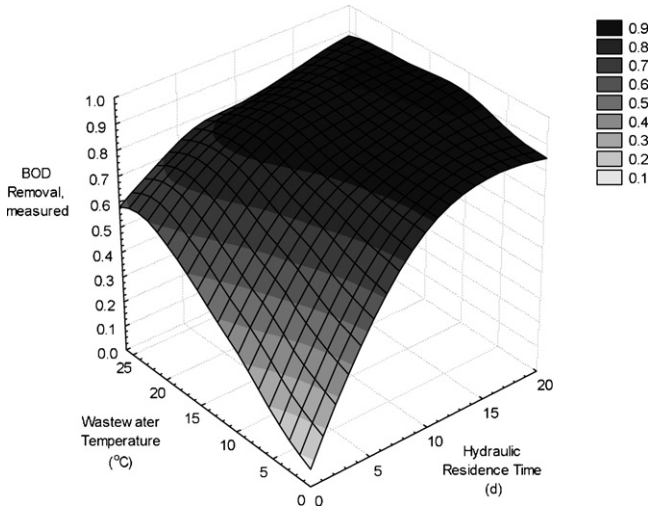

Fig. 2. Fitting surface of BOD removal over HRT and wastewater temperature.

beginning to drop afterwards (probably because of evapotranspiration effects during the summer months, which concentrates effluents, and decrease of temperature in the fall) until it reaches its minimum at winter start. Seasonal effects are more intense for lower HRT values, which implies a greater effect of temperature on BOD removal for experiments with small HRTs; it will be shown in Section 4 below that this behavior is predicted by the design equation for BOD removal that is proposed in the present work. Furthermore, there is much scatter for low HRT values, possibly because the wetlands do not have enough time to smooth out large-scale disturbances (like intense rainfall) in this case.

From Fig. 1a and b it is concluded that major factors affecting BOD removal are HRT and wastewater inlet temperature; this agrees with earlier studies [11]. To get a first idea on the functional dependence of BOD removal on HRT and wastewater temperature, a fitting surface of all the 812 BOD removal data points of the present work was constructed in the 3D space over HRT and wastewater temperature, and is depicted in Fig. 2. Distance-weighted least square fitting was used to construct the surface (meaning that data points that are close to some (X, Y) value combinations are used to create a local second-order regression polynomial, but with influence, i.e., weight, inversely proportional to their distance from the (X, Y) values). Based on Fig. 2, HRT is identified as the most important factor for BOD removal. There is a striking abrupt increase of the performance with increasing HRT at low HRT values, which is most apparent at low temperatures. The performance appears to reach a "plateau" at HRT values around 6–8 days (this limit depends on temperature), with a slight increase towards 1 (100% BOD removal) for higher HRT values.

As a first step towards PCA, cross-correlation coefficients between every variable that was considered a candidate for PCA (e.g., wastewater temperature, HRT, month, porosity, D_{50}, plant type, porous material type, meteorological data, pH, conductivity and DO) were evaluated. The cross-correlation coefficients were calculated with the use of following Eq. (6) and the results are presented in Table 3:

$$r = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}} \quad (6)$$

where x_i and y_i the parameter values and $\bar{x}$ and $\bar{y}$ the mean values of the two parameters. Values of the cross-correlation coefficients range from −1 to +1, with −1 implying perfectly inversely correlated variables, 0 pointing to zero correlation between the two variables, and +1 implying perfectly correlated variables. Apparently −1 is equally important to +1 regarding the importance of some variable in PCA, thus the absolute value of the cross-correlation coefficient is used in the following discussions.

In Table 3, D_{50} (i.e., the substrate material diameter that corresponds to the 50% line on grain size curve) is divided by the difference of maximum and minimum grain diameters ($D_{max} - D_{min}$); "porosity" refers to the porosity of the unit and not to that of the substrate alone, that is, it also includes the volume occupied by plant roots; the variables labeled "Material" and "Plant" are "binned", i.e., they get integer values different for every material and plant, respectively (e.g., "Plant" is 0 for the unplanted wetland, 1 for *Typha* and 2 for *Phragmites*) and "Material" is the porous media type (e.g., 1 for quarry material and 2 for river bed material). From the cross-correlation coefficients presented in Table 3 the variables correlated strongly to removal are HRT (absolute value of $r = 0.55$), month (absolute value of $r = 0.30$) and wastewater temperature (absolute value of $r = 0.32$). These three are the wetland "operational parameters." For the other parameters, the r values are relatively low. Temperature and month have a relatively high cross-correlation coefficient to each other (absolute value of $r = 0.42$). HRT has relatively low cross-correlation coefficients with both temperature (absolute value of $r = 0.08$) and month (absolute value of $r = 0.28$), and is an independent parameter which can be used in model construction. A group of four parameters is identified, i.e., porosity, $D_{50}/(D_{max} - D_{min})$, material and plant called "wetland unit parameters", with high absolute values of the cross-correlation coefficients ranging from 0.50 to 0.93. These parameters have a strong relation to each other, but show small dependence on BOD removal (absolute values of the cross-correlation coefficients ranging between 0.02 and 0.07), so they are not used in the construction of any model. "Meteorological parameters" and "physicochemical parameters" (DO, pH and conductivity) present low absolute values of the cross-correlation coefficients with BOD removal ranging from 0.00 to 0.12, and they were not tested in PCA. The conclusions drawn from the above analysis are confirmed by the following PCA.

3.2. Principal components analysis

The results of PCA, that is, the "loading factors" of the candidate variables on the principal components axes (see Section 2.2 for details) are reported in Table 4. The variables in this table are the same as in Table 3. A number of three principal parameters were found to explain nearly 83% of the variance in the original dataset (this will be the number of the simple ANN input parameters in Section 3.3.1 below).

Table 3
Variable cross-correlation coefficient r matrix

Parameter	Operational parameters			Wetland unit parameters			Meteorological parameters					Physicochemical parameters			Removal
	Month	T	HRT	Porosity	$D_{50}/(D_{max}-D_{min})$	Material	Plant	Relative humidity	Pressure	Solar heat flux	Wind speed	Rainfall depth	Conductivity	pH	DO	
Month	1.00	−0.42	−0.28	−0.03	−0.01	0.01	0.00	0.08	0.24	−0.62	0.36	−0.13	0.00	0.00	0.47	−0.30
T	−0.42	1.00	0.08	−0.00	0.00	−0.00	−0.00	−0.18	0.07	0.76	−0.49	−0.13	0.00	0.00	0.38	0.32
HRT	−0.28	0.08	1.00	0.01	0.00	−0.00	−0.00	−0.04	−0.02	0.16	−0.17	0.28	0.00	0.00	0.01	0.55
Porosity	−0.03	−0.00	0.01	1.00	0.70	−0.93	−0.68	−0.01	−0.02	0.01	−0.02	−0.01	0.32	0.12	0.27	0.02
$D_{50}/(D_{max}-D_{min})$	−0.01	0.00	0.00	0.70	1.00	−0.81	−0.50	0.01	−0.01	0.00	−0.01	0.00	0.27	0.09	0.30	0.02
Material	0.01	−0.00	−0.00	−0.93	−0.81	1.00	0.61	−0.01	0.01	−0.01	0.01	−0.00	0.12	0.11	0.01	−0.03
Plant	0.00	−0.00	−0.00	−0.68	−0.50	0.61	1.00	−0.01	0.00	−0.00	0.01	−0.00	0.09	0.15	0.28	−0.07
Relative humidity	0.08	−0.18	−0.04	−0.01	0.01	−0.01	−0.01	1.00	−0.26	−0.56	−0.12	0.35	0.00	0.00	0.00	−0.05
Pressure	0.24	0.07	−0.02	−0.02	−0.01	0.01	0.00	−0.26	1.00	0.04	−0.02	−0.45	0.01	0.00	0.00	−0.02
Solar heat Flux	−0.62	0.76	0.16	0.01	0.00	−0.01	−0.00	−0.56	0.04	1.00	−0.38	−0.21	0.05	0.08	0.15	0.01
Wind speed	0.36	−0.49	−0.17	−0.02	−0.01	0.01	0.01	−0.12	−0.02	−0.38	1.00	0.14	0.00	0.00	0.00	−0.01
Rainfall depth	−0.13	−0.13	0.28	−0.01	0.00	−0.00	−0.00	0.35	−0.45	−0.21	0.14	1.00	0.00	0.00	0.00	0.12
Conductivity	0.00	0.00	0.00	0.32	0.27	0.12	0.09	0.00	0.01	0.05	0.00	0.00	1.00	0.13	0.07	0.00
pH	0.00	0.00	0.00	0.12	0.09	0.11	0.15	0.00	0.00	0.08	0.00	0.00	0.13	1.00	.07	0.00
DO	0.47	0.38	0.01	0.27	0.30	0.01	0.28	0.00	0.00	0.15	0.00	0.00	0.07	0.07	1.00	0.00
Removal	−0.30	0.32	0.55	0.02	0.02	−0.03	−0.07	−0.05	−0.02	0.01	−0.01	0.12	0.00	0.00	0.00	1.00

Table 4
PCA results

	Component 1	Component 2	Component 3
Month	0.013	−0.846	0.030
T	−0.004	0.728	−0.525
HRT	−0.006	0.492	0.829
Porosity	−0.958	0.000	0.005
$D_{50}/(D_{max}-D_{min})$	−0.947	−0.009	−0.003
Material type	0.969	0.007	0.003
Plant type	0.771	0.006	0.003

Note: Variance explained 82.9% in total (47.8%, 21.3% and 13.8% for components 1, 2, and 3, respectively). The respective eigenvalues were: 3.348, 1.489 and 0.964.

The first principal component is readily recognized as expressing the "wetland unit parameters" group identified in the cross-correlation analysis above. All variables in this group have large loading factors on the first principal component, implying that any of them (with the possible exception of "Plant") could be used as input variable for the ANN. Of these variables, porosity was selected as the ANN input, because it is the handiest one for design and the more universal one (meaning that it could be used at different studies with different plant/substrate combinations not limited to the plants and substrates of the present work). Of course, every possible alternate parameter was tested, and found not to change the ANN results.

Concerning the second principal component, both wastewater temperature and month have large loading factors, but temperature was selected as ANN input variable, because this is the most conventional choice; seasonal effects expressed by month were used as input to an alternate ANN, described in Section 3.3.2 below. The third principal component is obviously related to HRT. Accordingly, there will be three input variables to the simple ANN that is presented in Section 3.3.1, namely porosity, wastewater temperature and HRT.

3.3. Artificial neural network

Two ANNs are presented in this section. The first ANN has three input variables that were identified in the PCA above, and it will be shown to perform reasonably well on the data of the present work. The second ANN is an alternative that includes several additional variables, in an effort to capture the effects of meteorological parameters, like precipitation, which may also be significant.

3.3.1. ANN with three input variables

As described in Section 2.2 of this work, an ensemble of 50 MLPs was trained for BOD removal over porosity, wastewater temperature and HRT. All members of the ensemble share a common topology, but they have different connection weights. Three neurons were found to be adequate for the hidden layer of the ANN. A member of the ensemble is depicted in Fig. 3a, and ANN predictions are compared to measured BOD removals in Fig. 3b for all the 812 BOD removal data points of this work (the straight line corresponds to $y = x$, i.e., perfect agreement between predicted and measured BOD removals). The wide range of

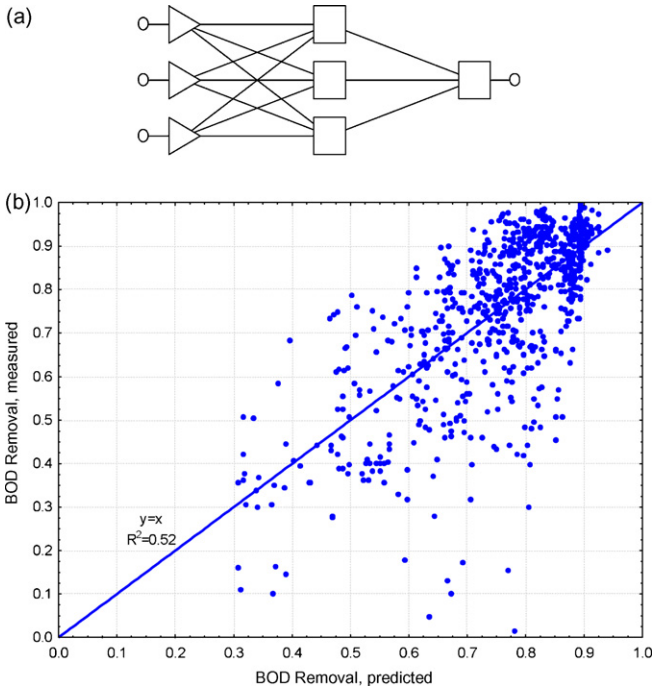

Fig. 3. (a) A member of the simple, 3-input-parameter ANN ensemble; (b) ensemble predictions over measured BOD removal.

removals (i.e., 0.3–1.0) is due to the wide range of HRTs used (i.e., 2–20 days). Although there is significant scatter around the straight line (the regression coefficient is relatively low for the data in Fig. 3b, $R^2 = 0.52$), the performance of the ANN is considered reasonably good for design of constructed wetlands. There are also some "outliers" (data points with exceptionally large difference between predicted and measured values) in Fig. 3b; most of them correspond to low HRT values and extreme wastewater temperatures (either very high, or very low temperature value), while some of the outliers also correspond to intense rainfall events.

The sensitivity analysis of the ANN response to changes of the input parameters confirmed the conclusions drawn from the study of the fitting surface in Fig. 2: ANN was found to be most sensitive to changes in HRT, while it was also sensitive to changes in wastewater temperature. The porosity of the wetland substrate was found to be of lower importance for the ANN predictions, at least for the porosities studied in the experiments of the present work. However, these porosities are typical of such systems ranging from 28% to 37%.

A useful result of ANN analysis is the "response curve" of the ANN to each one of the input variables. To construct the response curve for some input variable, the remaining input variables are assigned fixed typical values, and the network response to different values of the variable at hand are evaluated and graphed over the variable values. The response curve of the ANN discussed here with respect to HRT is shown in Fig. 4a. This ANN response curve is considered important, as it has a hyperbolic shape and combines zero and first order kinetics. The use of hyperbolic equations in describing the biological processes in wetlands has been proposed by Mitchell and McNevin [20] as an alternative to the most commonly used first-order decay law, and it is gaining acceptance thereafter. The response curve in Fig. 4a is fitted by a hyperbolic equation HRT/(HRT + 2), which combines zero and first order kinetics, while the temperature can be fitted by the function $(T^{-1} + 1.1)^{-1}$, with T again standing for wastewater temperature (Fig. 4b). The hyperbolic fits of Fig. 4 are shown in Section 4 below to be relevant to the design equation for BOD removal, which is proposed in the present work as an alternative to the ANNs.

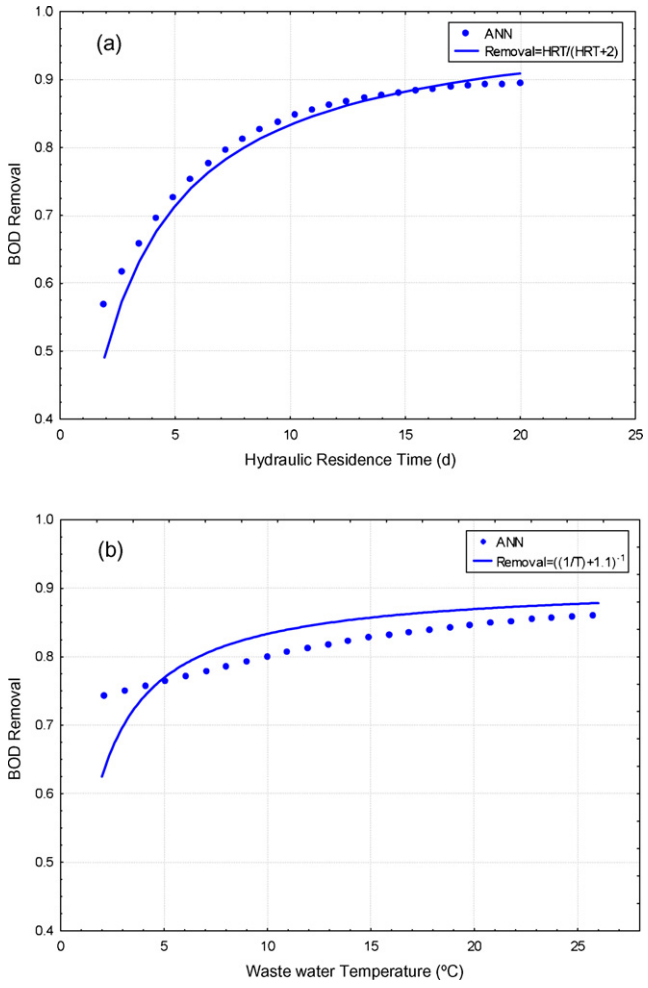

Fig. 4. ANN response curves with respect to (a) HRT; (b) wastewater temperature.

The method that was used to construct the ANN response curves can be used in pairs of ANN input variables as well, resulting to ANN response surfaces over combinations of two input variables. This is shown in Fig. 5 as the ANN response surface over HRT and wastewater temperature. This figure is comparable to Fig. 2 above, where a fitting surface of the raw BOD removal data over HRT and temperature is depicted. Comparison of the two figures confirms that the simple ANN presented here models BOD removal quite successfully.

3.3.2. An alternative ANN with more input parameters

Several ANN alternatives to the simple one presented above were studied in this work. Of those ANNs, the one with Month as an additional input variable to the three variables used above gave a slightly better agreement with experimental BOD

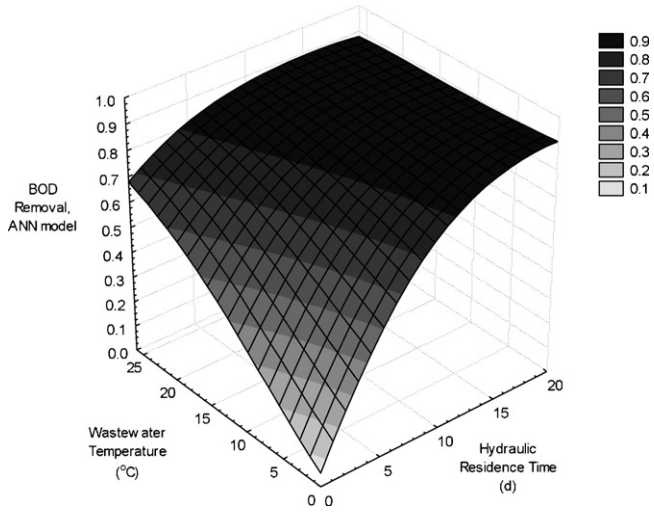

Fig. 5. ANN response surface over HRT and wastewater temperature.

removals, but it will not be presented here. Instead, a more complex ANN with the three input variables used in the previous section (i.e., HRT, wastewater temperature, and porosity) and six additional variables (i.e., month, and the following five meteorological variables: air humidity, barometric pressure, solar heat flux, wind speed and rainfall depth) will be discussed here in brief. The values of the meteorological parameters used here are averages over a time period of HRT days before measurement (e.g., for an HRT of 6 days, a 6-day average was used). Even though the meteorological parameters did not show a high cross-correlation with BOD removal (Table 3), the purpose of this ANN was to explore the possibility of improving (even slightly) modeling of the BOD removal process by inclusion of seasonal effects and large-scale hydrologic processes (that is, factors influencing evapotranspiration and precipitation).

In accordance to the network design principles of Section 2.2, an ensemble of 50 MLPs was trained over the experimental BOD removal values and the corresponding values of the 9 input variables. A number of six neurons in the hidden layer of the ANNs were found to be adequate in this case. Most of the results for this ensemble are similar to those for the simple network described earlier. ANN response was again found to be most sensitive on HRT and temperature changes; among the additional parameters of the new ANN, solar heat flux was found to be the more influential one (at high removal rates), and the ANN response curve with respect to this variable is shown in Fig. 6a, while the BOD removal predictions of the ANN are compared to measured BOD removal.

From the ANN response to solar heat flux, it is clear that BOD removal decreases with increasing evapotranspiration, as increased solar heat flux is related to increased evapotranspiration. Concerning the ANN performance, a comparison to Fig. 3b, where the predictions of the simple ANN are depicted, indicates improved modeling by the complex ANN. Although the regression coefficient for Fig. 6b ($R^2 = 0.68$) is again relatively low, it is higher than that for Fig. 3b ($R^2 = 0.52$) and the number of outliers for the complex ANN is clearly reduced. Therefore, it can be concluded that there is space for improvement of BOD

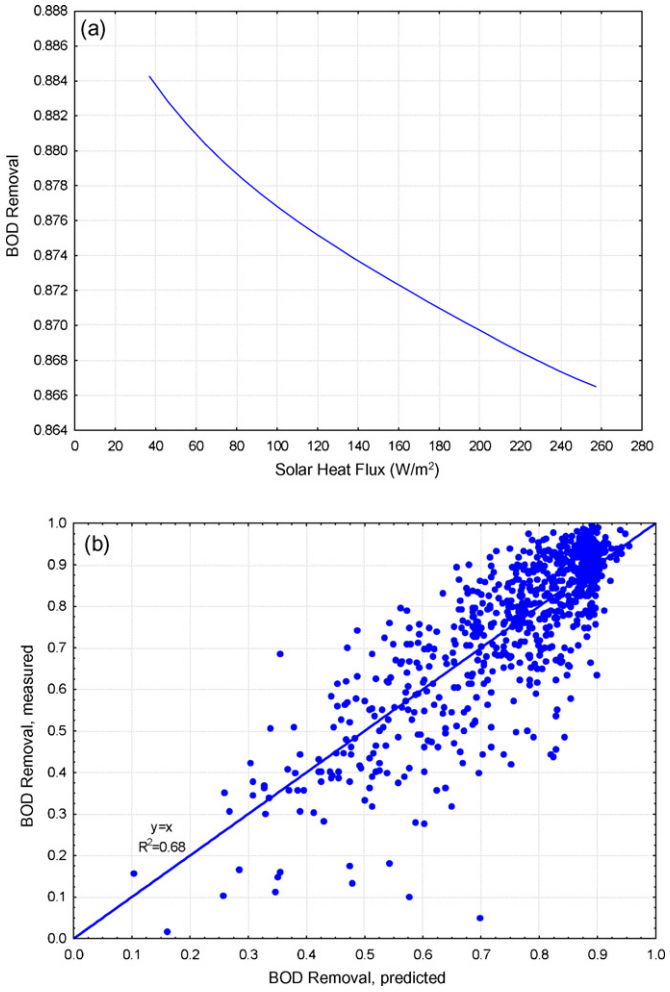

Fig. 6. (a) ANN response curve with respect to solar heat flux (W/m^2); (b) ANN predictions over measured BOD removals.

removal modeling by ANNs, if evapotranspiration is modeled and included together with precipitation and seasonal effects in the ANN input variables.

3.4. ANN validation

In an effort to test the ANN performance against other available data, a compilation of published BOD removal data was undertaken and is presented here; obviously, only the simple ANN can be tested, as there are no publications reporting the large number of input parameters of the complex ANN. Even so, several assumptions were required during the interpretation of the data. The data sources are presented in Table 5, along with the type of wetlands (which are all constructed wetlands of horizontal subsurface flow, but have various setups and treat different types of wastewater) and ranges of the inlet BOD, wastewater temperature, HRT and porosity for every data source.

All data reported in the data sources were used (82 data points in total). The data from Tanner et al. [21] include $CBOD_5$ values, as only these values were available. Mean annual temperature values were used for the data from Tanner et al. [21] and Ciria et al. [22], while a typical temperature of 15 °C was used for the

Table 5
Validation data description

Reference	Wetland use	BOD_{in} (mg/L)	Removal (%)	Temperature (°C)	HRT (d)	Porosity
[21]	Dairy farm wastewaters	57	60–92	15	2	0.37
[28]	Units downstream from stabilization ponds receiving domestic wastewater	25–100	72–82	25.5	6–12.5	0.40
[22]	Municipal raw wastewater	340–485	83–97	13.5	4.7	0.33
[23]	2.3:1 mixture of dairy parlor effluent and domestic sewage	451	89–94	15	10	0.30 and 0.33
[27]	Single household domestic wastewater	73–562	55–85	7–30	3.7–6.7	0.38 and 0.40
[24]	Primary settled municipal wastewater	113.7 – 131.1	54–94	15	1.7–16.1	0.40
[25]	14 systems treating municipal, industrial, domestic and hospital wastewaters	5–51	20–92	15	0.7–5	0.40
[8]	21 single-family domestic effluent systems	10.3–193.3	27–95	13	4.34–30.36	0.38
[28]	Pretreated swine effluent	343–411	86–92	22–25	4.3–14.7	0.37

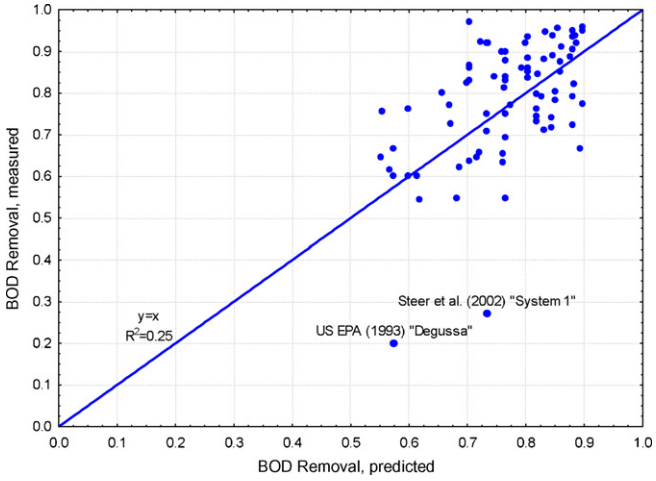

Fig. 7. ANN validation against BOD removal data from other studies.

data from Mantovi et al. [23] (Region Emilia, Italy), Headley et al. [24] (South Wales, Australia) and US EPA [25] (Los Angeles). Wetland porosities were judged based on the typical size of the material used as the wetland substrate, according to Reed et al. [26], for the data from Ciria et al. [22], Mantovi et al. [23], Karathanasis et al. [27], Steer et al. [8] and Lee et al. [28], while a typical porosity of 0.40 was used for the data from US EPA [25]. Inlet BOD_5 values were estimated from the wetland loading for the data from Senzia et al. [29] and Karathanasis et al. [27]. HRT values were estimated from the Hydraulic Loading Rate (HLR) for the data from Senzia et al. [29], and from the inlet flow rate and wetland dimensions for the data from Karathanasis et al. [27] and Steer et al. [8], with the inlet flow rate in the latter case estimated as 0.285 m^3/d/p.e., like in Karathanasis et al. [27]. Furthermore, both the BOD removal data and the values of the ANN input parameters in all data sources are averages over extended time periods, typically characterized by large standard deviations.

The ANN predictions are compared to the reported BOD removals in Fig. 7. Considering the gross assumptions made during the data compilation, the performance of the model is considered reasonably good ($R^2 = 0.25$). A linear regression equation was also fitted through the data of Fig. 7 ($y = 1.0173x$); its slope shows that the ANN slightly underestimates the measured data. There are two obvious outliers in Fig. 7, both of which correspond to units with exceptionally low influent BOD values, namely, US EPA [25] "Degussa" unit with inlet BOD = 5 mg/L and Steer et al. [8] "System 1" unit with inlet BOD = 10.3 mg/L. As reported in Table 5, these two inlet BOD values are the lower values in the entire dataset, and are in the range where BOD production within the wetland becomes important. It is known that ANNs in general cannot extrapolate outside the data range used for their training. The inlet BOD concentrations of the present work ranged from 282 to 507 mg/L, and HRTs ranged from 2 to 20 days. The network, however, performed well on data with inlet BOD concentrations as low as 20 mg/L, and with HRT values of 1 day; it was also able to handle units treating several types of wastes, although it was trained by data for synthetic wastewater. Further study of the applicability limits of the simple ANN presented in this section requires more data from controlled BOD removal experimental studies, particularly in real operating systems, but the results of the present work appear very promising.

4. A design equation for BOD removal

As mentioned before, the most commonly used equation (Eq. (2)) for BOD removal in subsurface flow constructed wetlands is the first-order equation. According to this approach, the hydrodynamic operation of the wetlands resembles that of plug-flow reactors (PFR), and there is a first-order decay of organic matter with time. Using the BOD removal definition from Eq. (5), Eq. (2) becomes:

$$1 - R = e^{-k_v t} \tag{7}$$

The rate constant at some temperature is estimated by an Arrhenius relationship (Eq. (3)). A designer wishing to evaluate HRT for some unit that will result to a BOD removal R_{BOD} is then faced with two constants with possible values spanning over extended ranges. The number of constants raises to three if the concept of "background concentration" C^* is used [4]. This is a portion of the C_{in} that is immediately settled down at the unit inlet, and cannot be removed; it is of the order 10–20 mg/L, and

is included in the formulation above by subtracting it from both the nominator and the denominator of the left hand side of Eq. (2).

$$\frac{C_{out} - C^*}{C_{in} - C^*} = e^{-k_v t} \quad (8)$$

An alternative to the first-order decay equation is the "Tanks-in-Series" model that predicts BOD removal based on the following equation [19]:

$$1 - R \equiv \frac{C_{out}}{C_{in}} = \frac{1}{(1 + k_v \cdot HRT/N)^N} \quad (9)$$

where k_v is again the volumetric rate given by Eq. (2), and N is the number of continuous stirred tank reactors (CSTR) in series, whose composite hydrodynamic behavior resembles that of the wetland. That is, this model considers a series of CSTR as more suitable than a PFR to describe the hydrodynamic behavior of a wetland. This approach is typically used for large-scale complex wetlands with non-trivial residence time distributions.

As shown in Table 1 of this work, typical values for k_{20} and θ in Eq. (7) are 1.104 d^{-1} and 1.06, respectively [25]. The predictions of the first-order decay model for the data of the present work were evaluated from Eqs. (2) and (8) above for these values of k_{20} and θ, and are checked against measured BOD removals in Fig. 8. It is obvious that the first-order model with these typical values of k_{20} and θ (Fig. 8a) overpredicts the present data ($y = 0.8x$, $R^2 = 0.33$). A fitting procedure suggests that k_{20} should be set to 0.88 d^{-1} (which is within the range of the values reported for this parameter in Table 1), θ should be set to 1.07 and C^* to 75.48 mg/L (Fig. 8b). The agreement with measured BOD removal becomes better but it is still not satisfactory ($y = x$, $R^2 = 0.34$). Furthermore, the C^* value is very high. The performance of the first-order model does not become much better when trying different values for θ. The limitations of the first-order model become apparent at this point. Besides, the fitting procedure, even if successful, would not shed too much light to the operation dynamics of the wetlands.

A design equation for BOD removal is proposed then in this work, as an alternative to the first-order model, whose shortcomings became obvious above. This equation, which follows, is based on the ANN response curves for HRT and temperature presented in Section 3.3.1:

$$R_{BOD} = \frac{HRT}{K + HRT} \quad (10)$$

where K has units of time (d), and is given by $K = c/T$ with c (d °C) a constant and T the wastewater temperature. The parameter K of Eq. (10) represents a time scale of the degradation process of organic matter. For the data of this work, and K and HRT expressed in days, a fitting procedure resulted to a value of 22.8 d °C for constant c. Thus:

$$R_{BOD} = \frac{HRT}{(22.8/T) + HRT} \quad (11)$$

The hyperbolic equation above is based on HRT, as this is considered most handy for design of wetlands. It will be shown below that Eq. (10) is an approximation of the first-order model,

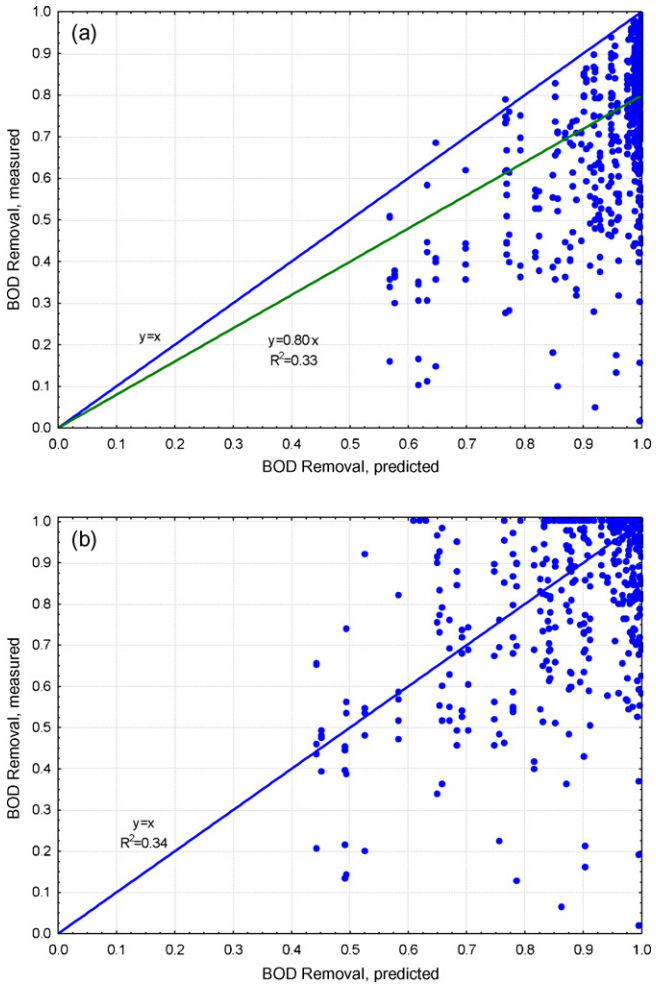

Fig. 8. Predictions of the first-order model compared to the BOD removal data of the present work.

thereby carrying the theoretical validity of this model, but also having a balance in the denominator of Eq. (10) between the degradation time scale K and HRT, with the former becoming critical for low HRT and losing importance for high HRT values (as seen in Figs. 2 and 5). Therefore, Eq. (11) describes more accurately, compared to the first-order model, BOD removal at low and a wider range of HRT values.

A fitting surface of the predictions of Eq. (11) for the BOD removal data of this work over HRT and wastewater temperature is shown in Fig. 9a. This figure is comparable to Fig. 2. A comparison of the two figures indicates that the design equation proposed here does capture the complex interplay between HRT and temperature for the BOD removal process. The predictions of Eq. (11) for the BOD removal data of this work are depicted in Fig. 9b against measured values (once again, the straight line in Fig. 9b is the "perfect agreement line" $y = x$, $R^2 = 0.44$). The performance of the design equation is similar to that of the simple ANN shown in Fig. 3b, and the discussion presented there applies here as well.

To express Eq. (10) as an approximation of the first-order model, the following classic equality of mathematics for the

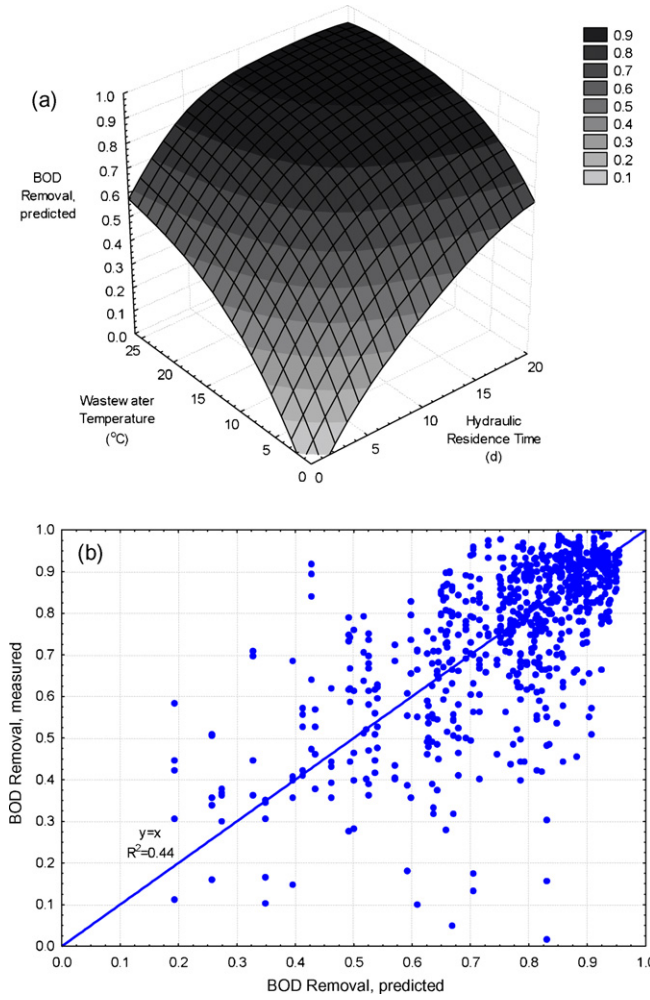

Fig. 9. (a) Fitting surface of BOD removal predictions of Eq. (11) for the data of this work over HRT and wastewater temperature; (b) predictions of Eq. (11) over measured BOD removal.

exponential function is used:

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} \tag{12}$$

Keeping only the first two terms of the right hand side of Eq. (12) and applying it to Eq. (7) one gets:

$$1 - R_{\mathrm{BOD}} = \frac{1}{1 + k_{\mathrm{v}} \times \mathrm{HRT}} \tag{13}$$

Interestingly, this result is the one given from Eq. (9) above as well, for the simpler case of the Tanks-in-Series model with $N = 1$. Eq. (13) can also result to a form similar to Eq. (10):

$$R_{\mathrm{BOD}} = 1 - \frac{1}{1 + k_{\mathrm{v}} \times \mathrm{HRT}} = \frac{1 + k_{\mathrm{v}} \times \mathrm{HRT} - 1}{1 + k_{\mathrm{v}} \times \mathrm{HRT}} = \frac{k_{\mathrm{v}} \times \mathrm{HRT}}{1 + k_{\mathrm{v}} \times \mathrm{HRT}} = \frac{\mathrm{HRT}}{(1/k_{\mathrm{v}}) + \mathrm{HRT}} \tag{14}$$

Accordingly, the parameter K in the design Eq. (10) is recognized here as an approximation for $1/k_{\mathrm{v}}$, with k_{v} the volumetric rate constant of the degradation process given by Eq. (2) above.

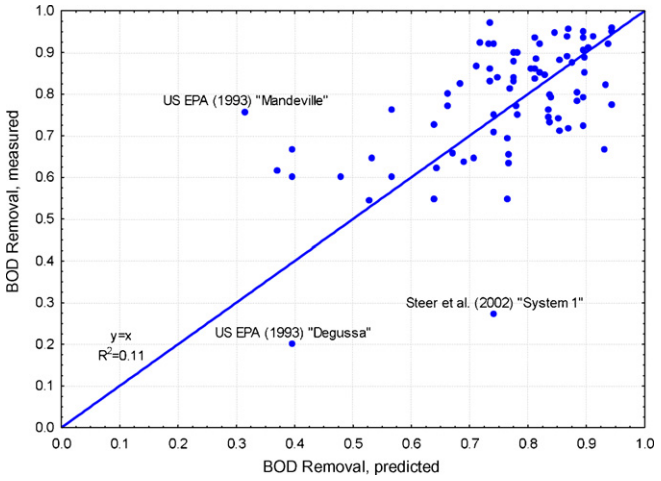

Fig. 10. Validation of the design equation for BOD removal against data from other studies.

The design equation proposed in this work is then a simple and more flexible, single-constant variant of the first-order model for BOD removal (the flexibility of the equation stems from the narrow range of possible values anticipated for the equation constant). The reason for the success of this approximating expression should probably be looked for at the smoothing that is inherent to biological processes, as compared to chemical reactors: this smoothing could drop out the higher order terms of the exponential decay.

Another feature of the design equation proposed here, that could be convenient in practical applications, is that the equation can be linearized. If both sides of Eq. (10) are inversed, one gets:

$$\frac{1}{R_{\mathrm{BOD}}} = \frac{K}{\mathrm{HRT}} + 1 = \frac{c}{T \times \mathrm{HRT}} + 1 \tag{15}$$

and the equation above could be used for estimation of c or for validity tests of the design equation, provided that experimental data for R_{BOD}, spanning a reasonably wide range, are available (obviously, Eq. (15) should not be used if all the data for R_{BOD} are in a narrow range, e.g., between 0.9 and 1.0).

The design equation (with the constant $c = 22.8$) is compared in Fig. 10 against the published dataset for BOD removal compiled in Section 3.4. Again, the straight line in this figure is defined by $y = x$. The performance of the equation is similar to that of the simple ANN of this work depicted in Fig. 7, and it is considered reasonably good for design purposes. The two outliers already mentioned in the discussion of Fig. 7 also appear in Fig. 10. Furthermore, low HRT values impose another limit to the applicability range of Eq. (11). There is a group of three units from US EPA [25], namely Mandeville, Greenleaves, and Monterey, that correspond to both low $\mathrm{BOD_{in}}$ values (41, 36, and 39 mg/L, respectively) and low HRT values (0.7, 1.0, and 0.9 days, respectively) for which Eq. (7) was shown to underpredict BOD removal. It is concluded that use of the design equation proposed here is not recommended for systems with $\mathrm{BOD_{in}}$ values lower than 50 mg/L, especially when the HRT of the units is lower than 2 days.

On the high side of the $\mathrm{BOD_{in}}$ scale, Eq. (11) predicts the performance of the heavy-loaded units reported by Lee et al.

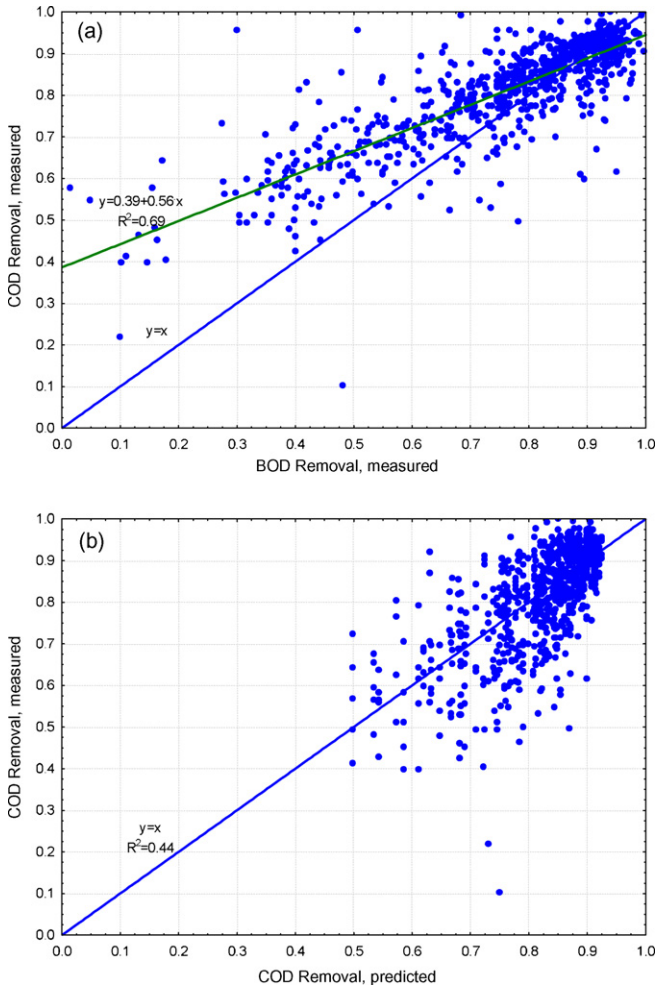

Fig. 11. (a) Correlation between measured BOD and COD removal; (b) predictions of the design equation proposed in this work for COD removal.

[28] very well. Although the present work does not distinguish between physical and microbial processes for BOD removal, Lee et al. [28] found the majority of BOD removal in their heavy-loaded units to be attributed to physical processes. It is then possible that the design equation proposed here (and the simple ANN of Section 3.3.1 as well) can predict BOD removal by both physical and microbial processes.

5. COD removal

The two parameters that express the removal of organic matter in constructed wetlands, i.e., BOD and COD, are widely considered to be correlated with equation COD = 1.8BOD [30]. For this reason, most researchers consider only one of these parameters in their studies. This is confirmed to be the case in the present work as well. Measured BOD removals are depicted against the corresponding measured COD removals in Fig. 11a (812 data points), with different symbols corresponding to different units and HRT values. A linear fit is drawn, in addition to the standard $y = x$ line (see for example Tanner et al. [21] for a similar correlation between COD and BOD), which has the following equation:

$$R_{COD} = 0.56\ R_{BOD} + 0.39 \quad (16)$$

Given the correlation above ($R^2 = 0.69$), COD removal could be predicted either by retraining the ANNs with COD removal data, or by modifying Eq. (11) for COD removal prediction, or even by applying the correlation between BOD and COD removal to the predictions of BOD removal. All these approaches gave results similar to each other and to the BOD removal predictions reported earlier in this work. Therefore, only the predictions of an equation similar to Eq. (10) are presented here in Fig. 11b (with the straight line again defined by $y = x$). A value of $c = 15.0$ was found to be the best for COD removal:

$$R_{COD} = \frac{HRT}{(15.0/T) + HRT} \quad (17)$$

The performance of this design equation appears to be reasonably good for COD removal, in spite of the low regression coefficient ($R^2 = 0.44$, like in every similar graph of this work). In particular, the number of outliers is considerably reduced in Fig. 11b.

6. Conclusions

Artificial neural networks have been shown in this work to be able to model the BOD removal process in horizontal subsurface flow constructed wetlands. Topologies of successful networks were suggested, and the network predictions were validated against an extended dataset reported here and a separate dataset compiled from studies published earlier. The performance of the networks was found to be reasonably good for wetland design purposes.

Based on the results of the artificial neural network modeling, a single-constant design equation was proposed here as a simple and flexible alternative to the first-order model commonly used for BOD removal prediction. The design equation was found to predict both datasets examined in this work reasonably well, and it was shown to be a hyberbolic mathematical approximation of the first-order model.

The COD removal in horizontal subsurface flow constructed wetlands was shown to be strongly correlated to the BOD removal, and to be predicted with adequate accuracy by both the neural networks studied here and the design equation proposed in this work.

Acknowledgements

The first author gratefully acknowledges a graduate studies scholarship from the Mpodosaki Foundation and from REVOIL S.A., Greece.

References

[1] P. Kuschk, A. Wiebner, U. Kappelmeyer, E. Weibbrodt, M. Kastner, U. Stottmeister, Annual cycle of nitrogen removal by a pilot-scale subsurface horizontal flow constructed wetland under moderate climate, Water Res. 37 (2003) 4236–4242.

[2] D. Hammer, Constructed Wetlands for Wastewater Treatment: Municipal, Industrial and Agricultural, Lewis Publishers, 1989.
[3] S. Reed, E. Middlebrooks, R. Crites, Natural Systems for Waste Management and Treatment, McGraw Hill, 1995.
[4] R. Kadlec, R. Knight, Treatment Wetlands, CRC Press, 1996.
[5] Q. He, K. Mankin, Performance variations of COD and nitrogen removal by vegetated submerged bed wetlands, J. Am. Water Resour. Assoc. 38 (2002) 1679–1689.
[6] A. Al-Omari, M. Fayyad, Treatment of domestic wastewater by subsurface flow constructed wetlands in Jordan, Desalination 155 (2003) 27–39.
[7] M. Greenway, A. Woolley, Constructed wetlands in Queensland: Perfomance efficiency and nutrient bioaccumulation, Ecol. Eng. 12 (1999) 39–55.
[8] D. Steer, L. Fraser, J. Boddy, B. Seibert, Efficiency of small constructed wetlands for subsurface treatment of single-family domestic effluent, Ecol. Eng. 18 (2002) 429–440.
[9] J. Vymazal, The use of sub-surface constructed wetlands for wastewater treatment in the Czech Republic: 10 years experience, Ecol. Eng. 18 (2002) 633–646.
[10] H. Brix, Constructed wetlands for municipal wastewater treatment in Europe, in: W.J. Mitsch (Ed.), Global Wetlands: Old World and New, Elsevier, Amsterdam, 1994, pp. 325–333.
[11] D.P.L. Rousseau, P.A. Vanrolleghem, N.D. Pauw, Model-based design of horizontal subsurface flow constructed treatment wetlands: a review, Water Res. 38 (2004) 1484–1493.
[12] S. Ould Dedah, W. Wiseman Jr., R.F. Shaw, Spatial and temporal trends of sea surface temperature in the Northwest African Region, Oceanol. Acta 22 (1999) 265–279.
[13] StatSoft, Inc. STATISTICA (data analysis software system), (2004) version 7. http://www.statsoft.com.
[14] Q. Chen, A.E. Mynett, Integration of data mining techniques and heuristic knowledge in fuzzy logic modelling of eutrophication in Taihu Lake, Ecol. Model. 162 (2003) 55–67.
[15] P.C. Nayak, Y.R. Satyaji Rao, K.P. Sudheer, Groundwater level forecasting in a shallow aquifer using artificial neural network approach, Water Resour. Manage. 20 (2006) 77–90.
[16] ASCE Task Committee, Artificial neural networks in hydrology-I: Preliminary concepts, J. Hydrol. Eng., ASCE 5 (2000) 115–123.
[17] ASCE Task Committee, Artificial neural networks in hydrology-II: Hydrologic applications, J. Hydrol. Eng., ASCE 5 (2000) 124–137.
[18] C.S. Akratos, V.A. Tsihrintzis, Effect of temperature, HRT, vegetation and porous media on removal efficiency of pilot-scale horizontal subsurface flow constructed wetlands, Ecol. Eng. 29 (2007) 173–191, doi:10.1016/j.ecoleng.2006.06.013.
[19] R.H. Kadlec, R. Axler, B. McCarthy, J. Henneck, Subsurface treatment wetlands in the cold climate of Minnesota, in: U. Mander, P.D. Jenssen (Eds.), Constructed Wetlands for Wastewater Treatment in Cold Climates, WIT Press, Southampton, Boston, 2003, pp. 19–52.
[20] C. Mitchell, D. McNevin, Alternative analysis of BOD removal in subsurface flow constructed wetlands employing Monod kinetics, Water Res. 5 (2001) 1295–1303.
[21] C.C. Tanner, J.S. Clayton, M.P. Upsedell, Effect of loading rate and planting on treatment of dairy farm wastewaters in constructed wetlands – I. Removal of oxygen demand, suspended solids and faecal coliforms, Water Res. 29 (1995) 17–26.
[22] M.P. Ciria, M.L. Solano, P. Soriano, Role of macrophyte Typha latifolia in a constructed wetland for wastewater treatment and assessment of its potential as a biomass fuel, Biosystems Eng. 92 (2005) 535–544.
[23] P. Mantovi, M. Marmiroli, E. Maestri, S. Tagliarini, S. Piccinini, V. Marmiroli, Application of a horizontal subsurfase flow constructed wetland on treatment of dairy parlor wastewater, Bioresour. Technol. 88 (2003) 85–94.
[24] T.R. Headley, E. Herity, L. Davison, Treatment at different depths and vertical mixing within a 1-m deep horizontal subsurface-flow wetland, Ecol. Eng. 25 (2005) 567–582.
[25] EPA, Subsurface flow constructed wetlands for wastewater treatment: A technology assessment, EPA, Washington D.C, 1993.
[26] S.C. Reed, D. Brown, Subsurface flow wetlands-a performance evaluation, Water Environ. Res. 67 (1995) 244–248.
[27] A.D. Karathanasis, C.L. Potter, M.S. Cayne, Vegetation effects on faecal bacteria, BOD and suspended solids removal in constructed wetlands treating domestic wastewater, Ecol. Eng. 20 (2003) 157–169.
[28] C.Y. Lee, C.C. Lee, F.Y. Lee, S.K. Tseng, C.J. Lia, Performance of subsurface flow constructed wetland taking pretreated swine effluent under heavy loads, Bioresour. Technol. 92 (2004) 173–179.
[29] M.A. Senzia, D.A. Mashauri, A.W. Mayo, Suitability of constructed wetlands and waste stabilization ponds in wastewater treatment: nitrogen transformation and removal, Phys. Chem. Earth 28 (2003) 1117–1124.
[30] Metcalf & Eddy, Inc., in: G. Tchobanoglous, F.L. Burton, D.H. Stensel (Eds.), Wastewater Engineering: Treatment and Reuse, fourth ed., MacGraw-Hill, New York, 2003.
[31] H. Brix, Use of constructed wetlands in water pollution control: historical development, present status and future perspectives, Water Sci. Technol. 30 (1994) 209–223.
[32] R.L. Knight, R. Ruble, R.H. Kadlec, S. Reed, North American treatment wetland database-electronic database created for the US Environmental Protection Agency, 1993.
[33] D.M.J. Griffin, R.R. Bhattairai, H. Xiang, The effect of temperature of biochemical oxygen demand removal in a subsurface flow wetland, Water Environ. Res. 4 (1999) 475–482.
[34] J. Vymazal, Czech constructed wetlands database, Ecology and use of wetlands, Prague, Czech Republic, 1998.
[35] A. Wood, Constructed wetlands in water pollution control: fundamentals to their understanding, Water Sci. Technol. 32 (1995) 21–29.
[36] R.H. Kadlec, Deterministic and stochastic aspects of constructed wetland performance and design, Water Sci. Technol. 35 (1997) 149–156.
[37] J. Vymazal, H. Brix, P.F. Cooper, R. Haberl, R. Perfler, J. Laber, Removal mechanisms and types of constructed wetlands, in: J. Vymazal, H. Brix, P.F. Cooper, M.B. Green, R. Haberl (Eds.), Constructed Wetlands for Wastewater Treatment in Europe, Backhuys Publishers, Leiden, 1998, pp. 17–66.
[38] H. Brix, Denmark, in: J. Vymazal, H. Brix, P.F. Cooper, M.B. Green, R. Haberl (Eds.), Constructed wetlands for wastewater treatment in Europe, Backhuys Publishers, Leiden, 1998, pp. 123–152.
[39] H.H. Schierup, H. Brix, B. Lorenzen, Wastewater treatment in constructed reed beds in Denmark-state of the art, in: P.F. Cooper, B.C. Findlater (Eds.), Proceedings of the International Conference on the use of constructed wetlands in water pollution control, Cambridge, UK, 24–28 September, 1990, pp. 495–504.
[40] P.F. Cooper, European design and operation guidelines for reed bed treatment systems, Prepared by EC/EWPCA Emergent Hydrophyte Treatment Systems Expert Contact Group, Water Research Centre, Swindon, UK, 1990.
[41] R.H. Kadlec, R.L. Knight, J. Vymazal, H. Brix, P. Cooper, R. Haberl, Constructed wetlands for pollution control: processes performance, design and operation, IWA specialist group on use of macrophytes in water pollution control, Sci. Tech. Rep. 8 (2000) 156.
[42] W. Liu, M.F. Dahab, R.Y. Surampalli, Subsurface flow constructed wetlands performance evaluation using area-based first-order kinetics, in: Proceedings of the Water Environmental Federation's Technical Exhibition and Conference (WEFTEC), Anaheim, CA, 14–18 October, 2002.