Indian Journal of Fundamental and Applied Life Sciences ISSN: 2231– 6345 (Online)
An Open Access, Online International Journal Available at www.cibtech.org/sp.ed/jls/2014/04/jls.htm
2014 Vol. 4 (S4), pp. 3732-3745/Pakrou et al.
Research Article

ARTIFICIAL NEURAL NETWORKS MODELING FOR PREDICTING TREATMENT EFFICIENCY AND CONSIDERING EFFECTS OF INPUT PARAMETERS IN PREDICTION ACCURACY: A CASE STUDY IN TABRIZ TREATMENT PLANT

***Saeid Pakrou[1], Naser Mehrdadi[2] and Akbar Baghvand[2]**
[1]Department of Environmental Engineering, Aras Campus for Tehran University, Iran
[2]Department of Environmental Engineering, Tehran University, Iran
**Author for Correspondence*

ABSTRACT
In this study, artificial neural networks modeling were evaluated for predicting the efficiency of Tabriz wastewater refinery that has activatedsludge treatment system with diffusion ventilation. Current activated sludge models involve characteristics of certain models and they have been constructed based on essential biosynthetic properties. However, it is very time-consuming and laborious to measure and calibrate these models. A computer model that is easily calibrated and that is user-friendly was used in this study, employing artificial intelligence techniques, i.e. artificial neural networks (ANN). These models cannot only be considered definite alternatives for existing models, but they can also connect to systems as error predictors. Using neural networks model and studies performed on the simulated hypothetic sludge refinery, constructed with the activated sludge process (SSSP) and its experiences, the system was modeled. For the hypothetic refinery, the produced simulation results were analyzed using the constructed neural network. In the hypothetic refinery's modeling, the highest correlation coefficient obtained with neural networks model was about 0.980 versus SSSP. Using real-life data from Tabriz wastewater refinery (collected over a six-month period from May to October 2013), the best and most appropriate neural networks model, constructed by combining input variables of Q_{inf}, TSS_{eff}, and MLss, resulted in the R value of 0.898 that shows a relatively high accuracy given the error ratio in the input variables.

Keywords: *Activated Sludge Process, Artificial Neural Network, Modeling, Efficiency, Tabriz Refinery*

INTRODUCTION
Although increasing growth of human societies and advances in industrial fields have led to special advantages, it has also caused numerous problems for the societies. One of these issues is the wastewater produced from residential places and from activities of industrial units (Arsiwala, 1993). Since improper disposal of urban and industrial sewage has undesirable effects on the environment, it becomes more important to refine wastewaters as completely as possible. If disposed of in the environment, urban wastewaters and more importantly industrial ones, lead to the pollution of surface waters and ground-waters, thus posing problems to reuse of the waters. Also, using water for different usages and severe water need in every district of Iran makes us consider preventing water waste in all ways and refine urban and industrial sewages, which are found in abundant volumes, thus trying to supply needed water.
These cases propound practical and control concepts in a quite complex manner. Operators of sewage treatment plants must be aware of information and technical, scientific, and operational knowledge in terms of their facilities as well as refinery equipment and processes so that they can have these operations under control and hence, qualify observation of conforming standards completely. In order to run operational routines connected with sewage treatment equipment in an optimal cost-effective way, it is necessary to apply the abovementioned skills completely, along with models of processing artificial neural networks.
In an urban sewage treatment plant, the three terms of primary treatment, secondary treatment, and advanced treatment are considered. The main goal in controlling water quality is to prevent spread of illnesses caused by water, and after being collected and conveyed to treatment plant, urban sewage is

refined during physical, chemical and biological operations. Leading and handling treatment plants are within specialty area of chemistry environment engineers (Shokouhi, 2008).

Refinery efficiency is highly affected by materials and properties of the input sewage. By default, every treatment plant has been designed for certain amount of input pollutants, temperature, etc. Variationsin any of these factors lead to deranged performance, thus demanding close attention of experts to keep efficiency.

The activated sludge system is a treatment system of aerobic suspending growth that was developed in 1914 by Ardern and Lockett and was made very popular in the same time because its process involved producing activated microorganisms that could also be active aerobically. Today, many instances of this process are used everywhere, but with similar basis (Tchobanglous and Bourton, 1991).

Existing models for the activated sludge model include ASM1, ASM2, ASM2d, and ASM3. All these models require to be calibrated for the certain systems on which they do refinery. In this regard, however, guiding systems such as artificial neural networks (ANNs) require less support.

Artificial neural network is a practical method for learning different functions such as functions with real values, functions with discrete values, and functions with vector values. Neural networks training is immune to the errors of training data, and these networks have successfully been applied to such problems as speech recognition, image recognition and interpretation, and robot training (Caudill, 1987). Artificial network is a method for calculation that is constructed based on joint connection of several processing units. The network is made from an arbitrary number of units or neurons that link the input set to the output. Among its features, calculation of a known function, an unknown function's estimation, pattern recognition, signal processing, and training could be mentioned. This method is very appropriate for problems where there are errors in training data, e.g. problems where training data have noise produced from sensors such as cameras and microphones, cases where samples are shown by a large number of feature-value pairs such as those from video cameras, there is enough time for learning and the target function has discrete values, there is enough time for learning (this method needs more time for learning compared to other methods e.g. decision tree) and there is no need for interpretation of target function (because it is hard to interpret weights learnt by the network) (Hamoda *et al.,* 1999). A type of neural network is made based on a computational unit called perceptron. A perceptron receives a vector of inputs and creates a linear mix of these inputs. If the result exceeds a threshold value, perceptron output will be equal to 1, -1 otherwise.

Using the method of modeling neural networks in sewage sector has so far been involving prediction of output wastewater quality. Accordingly, no significant caseshave been performed in connection with predicting the efficiency of wastewater treatment system using this model.

In a study based on artificial neural networks entitled *consideringfrequency effects of input parameters on prediction accuracy (COD) of industrial sewage using artificial neural networks,* Mehdipour and Shokuhian (2012) provided quality prediction of an industrial sewage treatment plant. In general, a total of 350tests have been used for experimentation and verification of the model. According to the results, chemical oxygen demand of input sewage (CODin), temperature, and pH are the factors that affect most the performance of a sewage treatment plant, in addition to chemical oxygen demand concentration.

In a study called *efficiency boost and economic strengthening of stabilization pond sewage treatment plants*, while expressing that one of the prospective issues of nationalwater and wastewater companies in the sector of utilizing wastewater treatment plants that use stabilization pond system is discharge and reuse of treatment plants, Kermani (2012) studied it. By suggesting that sludge leads to reduced pond efficiency, aerobic performance, and subsequently increase of output wastewater pollutants and reduced efficiency of treatment plants, he offered a practical cost-effective and feasible method.

In a study called *evaluation of treatment methods of wastewater treatment plants: Gheytarieh (Tehran) case study*, they considered increase of wastewater volume and infiltration efficiency. They found that Gheytarieh sewage treatment located in Northern Tehran, which currently refines about 145 m^3/h wastewater using activated sludge with wide depth ventilation, has become overloaded, and hence it requires development and improvement. The two methods of touch stabilization and oxygen ventilation have been compared in terms of economy, infiltration capacity increase, and utilization problems, and finally the best option was selected.

Given the importance of sanitary wastewater disposaland its proper treatment, Bagheri *et al.,* (2009) evaluated the ZanjanTreatment Plant in a study called *evaluating efficiency of Zanjan wastewater treatment plant*. The study was performed in a descriptive way via a cross-sectional study spanning for 10 years in 2008 (from May to February) in Zanjan treatment plant. Vector sample was observed on a daily basis on raw input wastewater and output wastewater, and efficiency of treatment plant was investigated by measuring parameters of TSS, BOD5, and COD. Considering that the efficiency of this treatment plant in removing sewage pollutants is about 84 %, the system used in this treatment plant was effective, and had reuse capability.

In a study called *evaluating statistical and artificial intelligence models for modeling wastewater treatment process,* Soltaninezhad *et al.,* (2012) considered this topic. Experimental data used for this study concerns daily measurements of the wastewater treatment plant in Minorca, Spain, recorded over 1990-1991. In the present study, results of wastewater treatment process in an activated sludge system of this city was modeled using a data-mining process and results were compared. Target problem was to find a suitable algorithmic model for predicting output wastewater quality considering input wastewater quality of the same day.

In a study called *optimization of dynamic systems of urban wastewater treatment plants,* Fili *et al.,* (2010) studied various factors and dynamic factors varying over time such as weathering, environmental, sanitary, maintenance, and technical factors via modeling, with purpose of efficient and appropriate utilization of an urban wastewater treatment system. By propounding different scenarios over time, they were optimized with the purpose of increasing efficiency and their economic justification. By identifying system bottlenecks in different time frames, operational schemes and scenarios were designed for removing bottlenecks.

In a study entitled *Pathology of Sewage Treatment Plants using the Activated Sludge Method with a Corrective Approach in Isfahan*, Mousavi *et al.,* (2009) suggested that the growing trend of the development and construction of sewage treatment plants in Iran has led experts toward optimization process, installing appropriate equipment, and making use of those equipment for observing standard wastewater output parameters. Results of investigations showed that improper use of pumps especially in sludge sector, impropriate choice of power panels, insufficient monitoring on picking electronic equipment according to corrosive environments of treatment plants, inefficiency of most measurement equipment, not sealing air openings, etc. have led to increased utilization and maintenance problems.

In a study called *Evaluating Industrial Treatment Plants using Artificial Neural Networks*, Mehdipour and Shokouhian (2012) considered the topic. Suggesting that improper functionality of sewage treatment might create serious problems for the environment and general health, they simulated performance of Mashhad Industrial Town's wastewater treatment plant using multilayer perceptron neural networks, which are among the most popular artificial neural networks used in environmental problems. Findings of the study showed that the provided neural networks model has an acceptable ability in predicting the performance of industrial treatment plants.

Observing that it is necessary to provide a reliable model for wastewater treatment plant with purpose of offering equipment for predicting and controlling processes to increase treatment plants' efficiency, Al-Asheh and Alfadala (2007) considered treatment plant behavior given input parameters in a neural network. When using neural networks, COD was suggested as an input for raw wastewater, with BOD and suspended solids being used in other circumstances, where results showed that better results will be obtained by considering all the three parameters.

Nasr *et al.,* (2012) considered Alexandria Wastewater Treatment Plant with purpose of minimizing utility cost and evaluating environmental balance stability in treatment plants and increasing its performance, concentrating on making use of the artificial neural networks (ANN). Such parameters as chemical oxygen demand (COD), biochemical oxygen demand (BOD), and total suspended solids (TSS), and data collected in studies over a one-year period were considered. This study implies that ANN can increase

plant performance with correlation coefficient of R between observed and predicted output variables to 0.90.
Raduly *et al.,* (2007) obtained treatment plants to evaluate performance and reliability via simulation in artificial neural networks, considering such factors as season temperature, rainfall amount, rainfall severity and duration, and effects of holidays on treatment plants, and they obtained similar results. Results show that artificial neural networks simulation offers acceptable results for wastewater system of an urban context, with an error less than 10%.
In present study, Tabriz Wastewater Treatment Plant was considered for 6 months for COD removal. Evaluated variables of the system included environment and wastewater temperature, flow rate, pH, turbidity, alkalinity, suspended solids (SS), COD, BOD5, and flow of small sewage and wastewater particles, in different units of treatment education. Data was obtained from the engineering department of the wastewater treatment plant.

MATERIALS AND METHODS

Tabriz Wastewater Treatment Plant is located 4 km west of Tabriz in the lands of GharaMolk, at southern corner of the river of Ajji Chai, and in the lowest spot of the city, in a land with the area of 72 hectares with a capacity of 765000 M^3/d, that uses activated sludge process. This wastewater treatment scheme has been planned for a population of 6 million people until 2025. It has also been designed in order to embody phosphorous and nitrogen removal units in the future.

Data

Daily data of $\theta_c, \Delta X, Q_{inf}, pH_{inf}, T_{inf}, COD_{inf}, MLSS, COD, and\ TSS_{eff}$ were collected over six months (from May to October 2013) from the control engineering department of Tabriz Treatment Plant, and necessary considerations were performed.

Modeling Neural Networks

Post-publishing Algorithm

During feed-forward process, each input unit (X_i) receives an input signal and forwards the signal to each hidden units of Z_1, Z_2,…, and Z_p. It then computes each hidden activation unit and forwards its signal (Z_J) to each output unit. Every activation output unit (Y_K) computes it as mere response for the given input pattern.
During training, each output unit compares its computed activation $Y_K$with its target value $T_K$to produce the error related to that sample for the unit. Accordingly, the factor of $\delta_K(k = 1,2, \ldots, m)$ has been computed. $\delta_K$is used for distributing error in the output unit of recursive $Y_K$across all units of previous layer. It is also used later for updating the weights between input and output hidden layers. In a similar way, the factor $\delta_j\,(j = 1,2, \ldots, p)$is calculated for each hidden unit Z_j.
Once all δ factors were calculated, weights of all layers are adjusted with simulation. Weight adjustment of W_{jk}(from hidden units of Z_j to output units of Y_k) is based on the $\delta_k$factor, and its activation Z_J from hidden unit of Z_j adjusts the weight V_{ij} (from input unit to X_i and hidden unit of Z_j) based on the factor δ_j and activation of X_i for the input unit.

Preprocessing

Preprocessing is needed for neural network weights. Picking weight impact has the effect that the network achieves the general state or an area in error reduction and in how it can converge. One epoch or a time period is a cycle according to training vectors or predefined tips. For a post-publishing training system of a neural network, many of these epochs are required.
The relationship between the number of available training patterns, number of weights that must be trained, and accuracy of intended classifications are denoted by P, W, and e respectively, that is, in short, shown in the form of W/P/e.
Data representation (data normalization) is very critical in neural networks. In many problems input and output vectors have a common part in a range in terms of value. In many practical applications neural networks can be supplied with variables with continuous value or a set or a range. The goal of a training network is to create balance between correct response to educational patterns and accurate responses to

new input patterns. This is equivalent to parameter estimation in definite traditional models. Keeping on training until total square errors are minimized is not always wise.

Hecht Nielsen suggests that the two different sets of data be used throughout training. One of these sets is used for weight adjustment, with the other being used in some time interval for error calculation. If error continues to reduce during the second experiment, the training is continuous. Once error starts to grow in the experiment set (called valuation in MATLAB'S Neural Network Toolbox), the network starts to retain the training patterns. Thus, this part of training will stop (Konar, 1999).

Halt criteria can be defined as criteria that lead the training to stop. Early stop is essential for generalization improvement. Existing data in this method are divided into the three following sets. The first subset is the training subset that has been used to compare error gradients and update network weight and orientations. The second subset is validation. Validation set errors is seen during the training process. Validation errors typically decrease in the first stage of training the same way that training set errors do. However, once the network starts to over fit data, errors in the validation set alsostart to increase naturally. When validation errors (called maximum validation failure in MATLAB's Neural Network Toolbox) for a certain number of repetitions increases, the training is stopped and weights and biases at the minimum of the validation error are returned (MATLAB Help, 2002).

Network structure had already been defined as the number of hidden neurons and hidden layers. Limitations concerning the number of neurons and hidden layers defined by Hecht and Nielsen (1987) and Roger and Dowla (1994) were selected as the basis for this study. Specifically, in neural networks models with a hidden layer, this upper limitation was used to generate MATLAB code. Some space problems and memory requirements exist in artificial neural networks with two hidden neurons.

Writing Code and MATLAB Neural Network Toolbox

Introduction to MATLAB Neural Network Toolbox and User Interface

The model was developed using the 2013 version of MATLAB software package from MathWorks Inc. MATLAB is a software package that consists of toolboxes for different engineering disciplines. Neural Network Toolbox (NNTool) is one of MATLAB toolboxes that implements artificial neural networks and is used in modeling process. To automatically search for a solution, a code has been writtenthat generates an artificial neural network for the function trained in an area predefined with a couple of neurons in the hidden layer, etc. The MATLAB toolbox consists of two graphical user interfaces (GUI). Network or information management is performed via connection between MATLAB console and Neural Network Toolbox as well as creating, manipulating, and adding or subtracting data in neural networks. The network offers an interface for initial value-assignment, training, simulation, etc. of artificial neural network in managing the network or information.

The graphical user interface can be divided into two sections of data and network representations on the top (input, output, targets, networks, errors, input delay of sections, and layer delay of sections). Data used can be extracted from the MATLAB console or MATLAB data files (with *.mat* extension). Entering, extracting, or removing new information or networks, and creating new networks is done with relevant buttons.

Second part of GUI-related content concerning network training is initialization as mentioned earlier. This section is a tool for the process of model construction and implementation. In the second part of main GUI, the second GUI shows that initialization, simulation, training, and conformity are created in neural networks and transferred to the main GUI. Second GUI does everything in neural networks, e.g. showing specifications of intended neural network being trained, initialization, simulation, synchronization, and showing neural network weights.

Neural Network Toolbox was mainly used, in this study, to extract more accurate results from a pre-developed neural network using code written for automatic generation of an artificial neural network model. Neural Network Toolbox is a user-friendly graphical user interface that can be used to generate a small number of models in certain areas.

MATLAB Coding

To develop the model, a MATLAB code has been written that automatically creates an artificial neural network for the sewage treatment plant. The code does the same thing done by Neural Network Toolbox GUI. To see accuracy changes, the code was used when different parameters of the generated neural network changed.

This code includes different versions in which some main structures of neural networks have been modified. For instance, a version of the written code constructs neural networks with only one hidden layer, and this code has three different versions in which the transfer functions have been modified. The same code behavior has been done with two hidden layers. The first version of the code has been included in the appendix. The code primarily receives a dataset from MATLAB Workspace. Next, the data is normalized in (0,1) and (-1,1) intervals using code *norm01.m* that has been prewritten in the interval (1,0) for range of [1,-1]. Data is then divided into three subsets of training, validation, and test sets. The code enters a loop that creates neural networks for 13 MATLAB train functions. These functions are listed in table 1.

For each of these train functions the script enters a second loop where changes of hidden neurons have been defined already. Before creating the dimensional network, the code trains the network and simulates the network with given data, thus performing a regression analysis of the results obtained in the first place. Next, the dimensional network is developed with the abovementioned method being applied. The written code also records R regression analysis form (correlation coefficient) for greater values versus some predefined threshold values, observed COD form versus predicted values, and the time series form for observed values versus those predicted from COD.

Table 1: Training functions used to create artificial neural network using the code

Train function	Short description
$Trainb$	Batch training with training rules for weight and bias
$Trainbfg$	BFGS quasi-Newton back-propagation
$Trainbr$	Bayesian regulation back-propagation
$Traincgb$	Conjugate gradient back-propagation with Powell-Beale restarts
$Traincgf$	Conjugate gradient back-propagation with Fletcher-Reeves updates
$Traincgp$	Conjugate gradient back-propagation with Polak-Ribiere updates
$Traingd$	Gradient descent back-propagation
$Traingda$	Gradient descent with adaptive learning rate back-propagation
$Traingdm$	Gradient descent with momentum back-propagation
$Trainlm$	Levenberg-Marquardt back-propagation
$Trainoss$	One-step secant back-propagation
$Trainscg$	Scaled conjugate back-propagation

Simulating Treatment Processes

Simulation of single-sludge processes (SSSP) is because of the fact that a very robust simulator has been selected for the activated sludge process including nitrification and de-nitrification. Furthermore, another reason for choosing this simulator is that it has been used in many successful modeling studies (Sin, 2000).

The SSSP system is a user-friendly program for simulating biological changes occurring in activated sludge of a treatment plant, performing carbon oxidation simultaneously, nitrification, and de-nitrification. The term *process rate* that denotes biological model changes was developed by the International Association on Water Pollution Research and Control (IAWPRC). These rateterms have been incorporated into 12 mass balances for heterotrophic and autotrophic biomass, substrate solution, nitrogen nitrates, and other significant compounds in analyses of the activated sludge process. Using numerical methods, this program offers a solution for input connected with the two constant time-dependent conditions.

Treatment model has formed as a chain from one to about nine reactors completely mixed. The user of this scheme specifically sees the process, pond volume, speed of flow into each reactor, solid materials, solid retention time (SRT), kinetic parameters, concentration of all components in feed flow, and the time-dependent pattern where all flows and concentrations flow. Typical values from kinetic parameters have been usedfor simulation in case kinetic data are not entirely available. Interested reader can refer to SSSP manual for model assumptions, parameters, and equations. Sensitivity analysis is applied on the set of this data.

Dynamic solutions in SSSP are used to calculate a solution for equations of mass equilibrium equations of input concentration and flow that changes in a 24-hour cycle. To use this typical method, input changes (flow pattern) must be provided in a 24-hour period. The SSSP automatically computes average flow rate and average concentration of input flow to use it for starting point comparison with the purpose of typical numerical integration. Then, the computer repeats integration for an over-a-24-hour cycle until change in components concentration is transferred from one cycle to the next. Put alternatively, the computer calculates answer of the system that receives the 24-hour input cycle every day. Therefore, the system response depends merely upon daily difference in loading (Bidstrup and Gradly, 1987).

Default dynamic datasets for the sample are provided as input file by the Typical Settled American Domestic Wastewater Constituent. On the cover of the installation package for this dataset is written "for heuristic simulation in a system's reaction to typical input changes". Average flow rate on the dynamic flow file is 1000 m^3/d (Bidstrup and Grady, 1987). This has been written in the SSSP manual that is a set of dynamic data offered solely for heuristic use.

Prediction accuracy is usually evaluated by offering that the network has not encountered before, known as ability, root mean square error (RMSE), and network generalization. For this purpose and to evaluate the designed networks, correlation coefficient criteria, mean absolute percentage error (MAPE), and mean absolute error (MAE) were used.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_{est} - Y_{act})^2}{n}} \quad (1)$$

$$R = \frac{\sum_{i=1}^{n}(Y_{act} - \bar{Y}_{act})(Y_{est} - \bar{Y}_{est})}{\sqrt{\sum_{i=1}^{n}(Y_{act} - \bar{Y}_{act})^2 (Y_{est} - \bar{Y}_{est})^2}} \quad (2)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_{est} - Y_{act}| \quad (3)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_{est} - Y_{act}}{Y_{act}}\right| \times 100 \quad (4)$$

Where, n: number of predictions, Y_{act}: actual values observed, Y_{est}:predicted values, $\bar{Y}_{act}$: average of actual values observed, and $\bar{Y}_{est}$: average of predicted values output from the model (Mehdi Pour and Shokouhian, 2012).

RESULTS AND DISCUSSION

Results

An activated sludge treatment unit was modeled in this study using the neural network method. One of the schemes was the hypothetical sewage treatment plant using version 1 of SSSP simulator and the other was the real treatment plant scheme in the city of Tabriz.

The hypothetical treatment plant has been constructed using default parameters of SSSP and default data for fine particles of a dynamic treatment plant representing a very simple treatment plant, and hence values for dependent values were obtained. Next, a fuzzy neural network model has been producedfor the real treatment plant using obtained data. The MATLAB code has been written for an automatic search process in error level with the purpose of minimum error determination in predictions from the trained model.

Studies for Treatment Process Simulation Model

Developing Neural Networks Models using SSSP Simulation Program

Development of neural network was done first with a hidden layer manually. Developed code was used to generate automatic model in stages after acquaintance with Neural Network Toolbox GUI. Using modeling user manual, two train functions were experimented (*trainb* and *trainbr*). *Trainb* trains the neural network model in batch mode that is one of the two weight updating modes. In this update mode, no update is performed as long as an epoch is not completed. This is the reason why in the neural network model, the training is done more slowly than other modes where updates are performed without waiting for the epoch to complete. The second training function used was trainbr in which weight and bias updates were typical given dedicated values in the Levenberg-Marquardt optimization method. Training function was Bayesian. Once code was developed for automatic network of neural network, this code was used for training. The code was straight-forward and worked as follows:

Number of hidden neurons for each training function varied from 2 to 10. 13 training functions were offered by MATLAB Neural Network Toolbox for implementing the back-propagation algorithm. For each training function, 9 neural network models were developed, given different number of hidden neurons from 2 to 10, and therefore 117 models for each complete execution cycle of the code. Neural network models created manually using Neural Network Toolbox GUI were much fewer than the number of models created by code.

Guide to Developing Neural Network Model using the Toolbox GUI

Manual application showed that the model created with 3 hidden neurons in the hidden layer had minimum error and the best proportion with actual data. During run, all other properties of the neural network were kept constant. Then the model with a hidden layer and 3 hidden neurons was selected as the optimal model, and hence training rate changed during training operation. Next, epoch size had a significant reduction to 6000, because no significant decrease was visible after this point in the mean squared error (MSE). Therefore, the application stopped at epoch 6000.

Properties of neural networks used in manual run have been listed in table 2. The input layer has three neurons, each one representing a parameter (X_s, X_i, Q) of the model. The output layer has three neurons each representing one variable to be predicted$(S_s, X_{het}, MLVSS)$.

Table 2: Properties for neural networks trials for each execution

Trial	training Function	Epochs	Goal	Learning Rate	max_fail	Learning Function
Trial1	Traingd	6000	0.05	0.070	5	*Learngd*
Trial2	Traingd	6000	0.05	0.080	5	*Learngd*
Trial3	Traingd	6000	0.05	0.085	5	*Learngd*
Trial4	Traingd	6000	0.05	0.090	5	*Learngd*
Trial5	Trainbr	6000	0.05	0.090	5	*Learngd*

Neural Network Model Data Studies for Tabriz Treatment Plant

Data Preparation for Tabriz Treatment Plant

Two different datasets were used for Tabriz Treatment Plant neural network model. Early executions related to the time of first use of datasets. However, the neural network model was developed using this dataset and weak predictions. The biggest correlation obtained was about 0.35. Afterward, the second dataset was used to develop models.

In the modeling process, 8 system variables were used to construct a neural network model. Also, combinations of different data in the development process have been used. System variable used during the modeling process included solid retention time (θ_c), flow speed of fine particles(Q_{inf}), pH for fine

particles, water temperature for fine particles (T_{inf}) , COD concentration of particles, MLSS, wastewater COD, TSS for wastewater, and the rate of sludge production from primary sedimentation tank.
These variables are used by themselves or in combination to predict wastewater COD concentration of treated water in treatment plants. Many combinations of variables were used and hundreds of models have been developed and trained to observe effectiveness of predictions. Only models with high effectiveness result are presented here.
Data preparation was done by removing values devoid of primary datasets. First, raw data consisted of daily measurements from about 6 months, except Fridays and 150 days for each test, equaling almost 450 data items totally. There were empty data items also. Data rows with empty values were completely removed from the set.

Table 3: Combining variables used in the process of model development

No	Q_{inf}	pH_{inf}	T_{inf}	COD_{inf}	MLSS	TSS_{eff}	ΔX	θ_c	Best R
1	✓	✓	✓	✓					0.411
2	✓	✓	✓						0.360
3	✓	✓	✓		✓				0.310
4	✓	✓	✓			✓			0.462
5		✓	✓	✓				✓	0.257
6		✓	✓					✓	0.223
7		✓	✓		✓			✓	0.207
8		✓	✓			✓		✓	0.211
9	✓	✓	✓	✓			✓		0.291
10	✓	✓	✓				✓		0.256
11	✓	✓	✓		✓		✓		0.302
12	✓	✓	✓			✓	✓		0.321
13	✓		✓	✓					0.451
14	✓		✓						0.421
15	✓		✓		✓				0.397
16	✓		✓			✓			0.459
17			✓	✓				✓	0.243
18			✓					✓	0.268
19			✓		✓			✓	0.267
20			✓			✓		✓	0.289
21	✓		✓	✓			✓		0.269
22	✓		✓				✓		0.244
23	✓		✓		✓		✓		0.201
24	✓		✓			✓	✓		0.231
25	✓				✓	✓			**0.898**
26	✓				✓	✓			**0.729**

Data Pre-processing of Tabriz Treatment Plant
After data preparation for model development, processing in weight update processwas done especially upon using non-linear transfer function. Data pre-processing was done by setting data input and output in the ranges of [0,1] or [-1,1]. This conversion was performed for all data points as described below:
For [0,1]:

$$X_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (5)$$

And for [-1,1]:

$$X_i = 2\times\frac{x_i - x_{min}}{x_{max} - x_{min}} - 1 \quad (6)$$

Each data variable must be fed using normalization of input data with non-linear transfer functions such as *tansig*, *logsig* with equal weight in the network model. Thus, data must be scaled in [0,1]or [-1,1] using equations (5) and (6).

While running with the range of [-1,1], no good result was obtained. In these executions, first the script was tested with a hidden layer. Afterwards, number of hidden neurons change using criteria defined by Hecht and Nielsen (1987). Highest correlation rate in these runs was about 0.4.

Next, number of hidden layers increased to 2. Using the same data and in the range of [0,1], *logsig* and *tansig*transfer functions were used respectively for ranges of [0,1] and [-1,1]. It is possible to use transfer function combinations in multilayer models using a certain transfer function in one layer and a different transfer function in another layer. Some possible combinations of these functions when using dual-layer hidden layers are as follows:$(tansig - tansig, logsig - logsig, tansig - purelin, \ldots)$.

Developing an Artificial Neural Network for Tabriz Treatment Plant

MATLAB software was used to construct the neural network model automatically using available data for Tabriz Treatment Plant. Variable combination used in the process of model development has been presented in table 3. These combinations were given in order, and attempts were made so that as long as a good fitting was obtained for data by visual judgment and given the correlation coefficient *R*, when an acceptable fitting was achieved, this has been enhanced to the best results by grouping variables.

As seen in table 3, the best fitting was achieved for wastewater COD data by combining 25 items for the correlation coefficient of R=0.898. Based on these observations, the variables specified in 25-item combination were mostly used alone or in couples to create new subsets in the script, and they were experimented in searching for improved results. Combination of the variable under examination is presented in table 4.

Table 4: Combination of variables experimented

Set No	Variable(s) used
1	MLSS only,
2	TSS_{eff} only,
3	MLSS and Q_{inf},
4	Q_{inf} and TSS_{eff},
5	Q_{inf}, TSS_{eff} and MLSS,
6	MLSS and TSS_{eff}.

In addition to the manual of combining variable combinations, the script automatically causes changes in the number of hidden neurons to improve output. Output here was to predict wastewater COD. All these runs have been performed for different hidden layers. In the first run, there was one hidden layer and the number of hidden neurons changed from 1 to 9. As explained earlier in the section for materials and methods, there are 13 training functions available in MATLAB Neural Network Toolbox for back-propagation algorithm. Using sequential functions, this script creates the neural network (table 1). Transfer functions *tansig* and *logsig* are run in different scripts. Regression graph produced by this script for tested models cannot have performance correlation coefficient exceeding 0.6. After that, the same method was used via script for artificial neural networks with two hidden layers. In subsequent runs using a script with two hidden layers, again the number of hidden neurons changes for both hidden layers.

The best fitting for neural network model was calculated according to highest correlation coefficient. Model neural network's response, regression analysis, and graphical session training in figures 2, 3, and 4, as shown in figure 2, first half of a graph including training data, includes validation and training data in the second half (AFTER X=37700). Results from regression analysis are shown in figure 3. Blue line

Indian Journal of Fundamental and Applied Life Sciences ISSN: 2231– 6345 (Online)
An Open Access, Online International Journal Available at www.cibtech.org/sp.ed/jls/2014/04/jls.htm
2014 Vol. 4 (S4), pp. 3732-3745/Pakrou et al.
Research Article

given on the curve *A=T*, represents the best fitting possible in visual terms and the red line shows the best fitting.

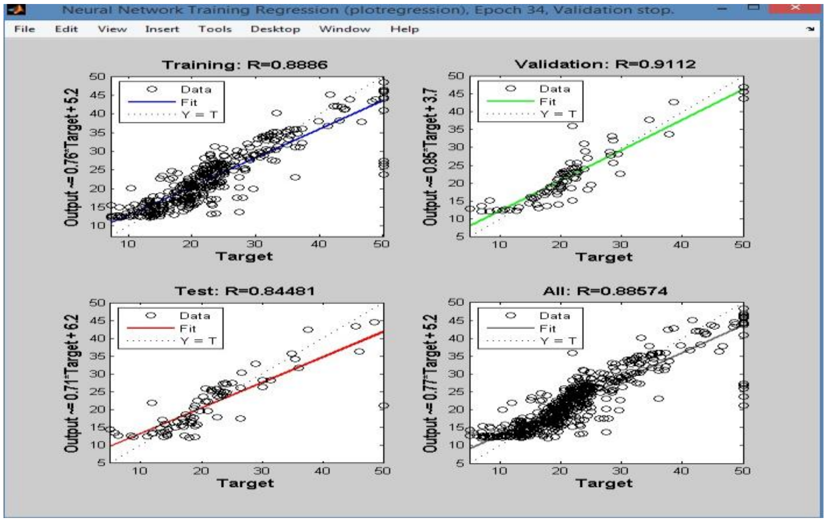

Figure 1: Regression output in neural network model

MSE (mean squared error) is given in figure 4 for the best performance in training session. Training session in epoch 48, was stopped after validation process finished. There are three rules for stopping training session: 1) achieving the goal that minimizes mean squared error (MSE) down to zero; 2) number of failures exceed a certain number; and 3) to achieve the maximum number of given epochs (here, it was 2000). In the best run, stop criteria of 2 for stopping the training session has been used.
The neural network model has been produced by running script 25 for Tabriz Treatment Plant (table 3). System variables used and run in this script include flow rate, TSS wastewater, and MLSS. 99 daily data items have been used in developing the model. Data were divided into three sections of training, validation, and test set. First 50 % of data (days 1-49) were considered as training data; 25% (days 50-74) as validation set used for validation with the purpose of error reduction in training (parameter estimation), and the remaining (days 75-99) as test data. Combination 26 was the same 25 with the difference that data were not fed in time order.

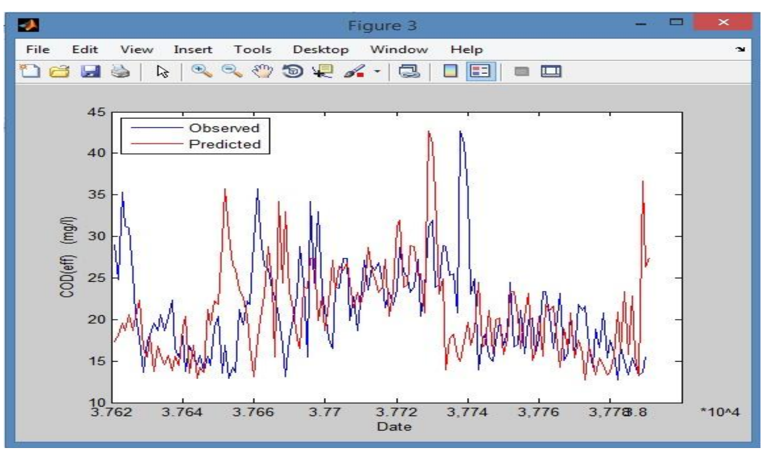

Figure 2: COD value prediction in the best neural network model (R=0.8978)

In previous figures it could be observed that obtained results were fairly good. The best correlation coefficient obtained was 0.8978. Results can even be better than different and accurate data used. Datasets used in this study had lots of void spaces, leading to the constant input interruption of the neural network

Research Article

model. Unreliable quality of daily data led to increased uncertainty. Although it is clear that a neural network can deal with noise and improper data, process parameters in biological wastewater treatment used in the neural network model has different and relative errors. Therefore, a combination of errors is published in the neural network model. Thus, the main ideal goal in the neural network model can make use of the data from automatic treatment plant sensors. Accordingly, accidental human error is minimized.

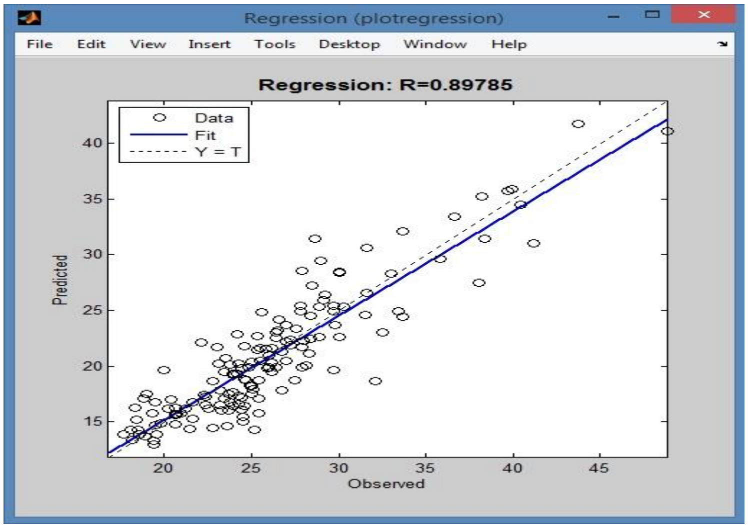

Figure 3: Regression analysis of the best neural network model

Best results obtained using MATLAB script is provided in tables 4-6. In this table, name of the back-propagation algorithm, variable number 74 of table 4 of the set, number of hidden neurons employed and correlation coefficient have been presented in total.

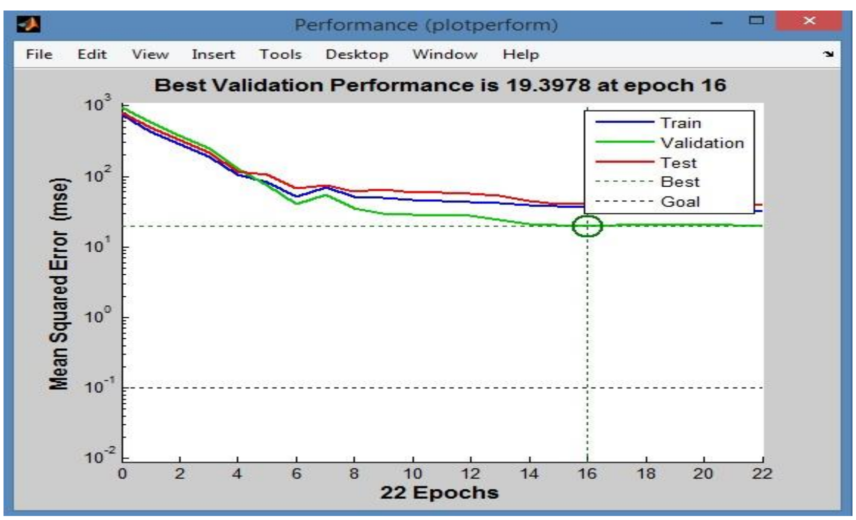

Figure 4: Prediction of mean COD error in the neural network model

The neural network model can be trained as a two-stage process. In the first stage, a combination of provided variables considered the best result for creating new variable sets. In this case given variables specified in set 25 in table 5 were employed, with new combinations exploited using these identified variables. Set 26 had been identified with difference in data order fed to the neural network model to 25. In set 26, data were fed randomly, whereas this was done in time sequence in 25.

Table 5: Best results in 25th run of code in MATLAB software

Training function	Set	HNs	R
Trainbfg	6	4	0.673
Traincgb	5	6	0.898
Traincgb	6	6	0.694
Trainlm	6	1	0.670
Trainscg	6	6	0.668

Conclusion

In this study, using neural network model and studies on simulation of the hypothetic treatment plant constructed with simulation of single-sludge processes (SSSP) and using experiences of that system, Tabriz Wastewater Treatment Plant was modeled in processing terms. For the hypothetic treatment plant, simulation results obtained by the developed neural network model were analyzed. In modeling the hypothetic treatment plant, the highest correlation coefficient obtained with neural networks model versus SSSP was 0.980.

Using actual data from Tabriz treatment plant, the best and most suitable neural network model achieved had an R value of 0.898 that shows a relatively high accuracy given existing error value in the input data.

REFERENCES

Al-Asheh S and Alfadala HE (2007). Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance, *Journal of Environmental Management* **83**(3) 329–338.

Arsiwala (1993). *Wastewater Treatment*, translated by Ahmad Reza Yazdanbakhsh and Kazem Nadaghi (Fardabeh Publishing).

Bagheri Ardebilian P, Nabi'ei A and BagheriArdebilian M (2009). Efficiency evaluation of Zanjan Wastewater Treatment Plant in 2008, *twelfth conference on environmental hygiene.*

Bidstrup SM and Grady Jr. CPL (1988). SSSP Simulation of single-sludge processes. *Journal of the Water Pollution Control Federation* **60** 351-361

Caudill M (1987). Neural networks primer: Part I, *AI Expert* 46-52.

Hamoda MF, Al-Gusain IA and Hassan AH (1999). Integrated wastewater treatment plant performance evaluation using artificial neural network, *Water Science and Technology* **40** 55–69.

Hecht-Nielsen R (1987). Kolmogorov's Mapping Neural Network Existence Theorem. *Proceedings of the First IEEE International Joint Conference on Neural Networks*, New York 11–14.

Kermani M (2012). Efficiency boost and economic strengthening of stabilization pond sewage treatment plants, *2nd conference on environmental management and planning*.

Konar A (1999). Artificial Intelligence and Soft Computing, Behavioral and Cognitive Modeling of the Human Brain.

Mehdi Pour Torghabeh A and Shokuhian M (2012). Considering performance of industrial treatment plants using artificial neural network model, *6th national conference and specialized exhibition of environmental engineering.*

Mehdi Pour Torghabeh A and Shokuhian M (2012). Considering the effect of input parameters on prediction accuracy of COD in wastewater of industrial treatment plants using artificial neural networks, *6th national conference on environmental engineering.*

Mousavi M, Farhadian M and Jaberi A (2009). Pathology of sewage treatment plants using the activated sludge method with a corrective approach in Isfahan, *3rd national conference on water and wastewater with usage approach.*

Nasr MS, Moustafa AE and Seif AE (2012). Application of Artificial Neural Network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT, *Alexandria Engineering Journal* **51**(1) 37–43.

Pasha Zanoosi S, Abbasi M, Aboutorab H and Ayati B (2008). Evaluating methods for wastewater treatment plants: Gheytarieh Wastewater Treatment Plant case study, *Second national conference on water and wastewater with usage approach.*

Phili H, Misavi Biouki H, Sakaki A and Dehghan Menshadi S (2010). Optimization of dynamic systems of urban wastewater treatment plants by simulation, *2nd national seminar on position of refined waters and wastewater in managing water resources.*

Ráduly B, Gernaey KV, Capodaglio AG, Mikkelsend PS and Henzed M (2007). Artificial neural networks for rapid WWTP performance evaluation: Methodology and case study, *Environmental Modelling & Software* **22**(8) 1208–1216.

Roger LL and Dowla FU (1994). Optimization of Groundwater Remediation using Artificial Neural Networks with Parallel Solute Transport Modelling. *Water Resources Research* **30**(2) 457-481.

Shokuhi R (2008). *Urban and Industrial Wastewater Treatment* (Mehrazan Publishing).

Sin G (2000). Determination of ASM1 Sensitive Parameters and Simulation Studies for Ankara Wastewater Treatment Plant. M.S. Thesis, METU.

SoltaniNezhad A, Afsahi M, Nakha'di Zadeh G and Chalkesh Amiri M (2012). Evaluating statistics and artificial intelligence models in modeling the wastewater treatment process, *National conference on water/wastewater engineering sciences.*

Tchobanoglous G and Burton FL (1991). Wastewater Engineering: Treatment, Disposal and Reuse. McGraw-Hill Series in Water Resources and Environmental Engineering