Effluent Quality Prediction of Wastewater Treatment System Based on Small-world ANN

Ruicheng Zhang
College of Electrical Engineering, Hebei United University, Hebei Tangshan 063009, China
Email: rchzhang@yahoo.com.cn

Xulei Hu
College of Electrical Engineering, Hebei United University, Hebei Tangshan 063009, China
Email: hxlei_2010@163.com

Abstract—In order to provide a tool for predicting wastewater treatment performance and form a basis for controlling the operation of the process, a NW multi-layer forward small world artificial neural networks soft sensing model is proposed for the waste water treatment processes. The input and output variables of the network model were determined according to the waste water treatment system. The multi-layer forward small world artificial neural networks model was built, and the hidden layer structure of the network model was studied. The results of model calculation show that the predicted value can better match measured value, playing an effect of simulating and predicting and be able to optimize the operation status. The establishment of the predicting model provides a simple and practical way for the operation and management in wastewater treatment plant, and has good research and engineering practical value.

Index Terms—NW small-world networks, multi-layer forward neural networks, wastewater plant, modeling, wastewater treatment

I. INTRODUCTION

The increased concern about environmental issues has encouraged specialists to focus their attention on the proper operation and control of wastewater treatment plants (WWTPs). The characteristics of influent to the WWTPs are varied from one plant to another depending on the type of community lifestyle. Therefore, the performance of any WWTP depends mainly on local experience of a process engineer who identifies certain states of the plant [1]. The type of influent for any plant is also time-dependent and it is difficult to have a homogeneous influent to a WWTP [2]. This may result in an operational risk impact on the plant. Serious environmental and public health problems may result from improper operation of a WWTP, as discharging contaminated effluent to a receiving water body can cause or spread various diseases to human beings. Accordingly, environmental regulations set restrictions on the quality of effluent that must be met by any WWTP.

A better control of a WWTP can be achieved by developing a robust mathematical tool for predicting the plant performance based on past observations of certain key parameters. However, modeling a WWTP is a difficult task due to the complexity of the treatment processes. The complex physical, biological and chemical processes involved in wastewater treatment process exhibit non-linear behaviors which are difficult to describe by linear mathematical models.

Owing to their high accuracy, adequacy and quite promising applications in engineering, artificial neural networks (ANNs) can be used for modeling such WWTP processes. The ANN can be used for better prediction of the process performance. It normally relies on representative historical data of the process. In a wastewater treatment plant, there are certain key parameters which can be used to assess the plant performance. These parameters could include biological oxygen demand (BOD), suspended solid (SS) and chemical oxygen demand (COD). Most of the available literature on the application of ANNs for modeling WWTPs utilized these parameters. For example, Oliveira-Esquerre et al. [3] obtained satisfactory predictions of the BOD in the output stream of a local biological wastewater treatment plant for the pulp and paper industry in Brazil. The principle component analysis was used to preprocess the data in the back propagated neural network. The Kohonen Self-Organizing Feature Maps (KSOFM) neural network was applied by Hong et al. to analyze the multidimensional process data and to diagnose the inter-relationship of the process variables in a real activated sludge process. The authors concluded that the KSOFM technique provides an effective analysis and diagnostic tool to understand the system behavior and to extract knowledge from multidimensional data of a large-scale WWTP. Hamed et al. [2, 4] developed ANN models to predict the performance of a WWTP based on past information. The authors concluded that the KSOFM technique provides an effective analysis and diagnostic tool to understand the system behavior and to extract knowledge from multidimensional data of a large-scale WWTP. The authors found that the ANN-based models provide an efficient and robust tool in predicting WWTP performance. But the large learning assignment, slow convergence, and local minimum in the neural network are observed.

For the problems in those wastewater treatment models mentioned above, in this study, a new wastewater treatment model, NW multilayer feedforward small-
world artificial neural network model, is proposed, which integrates NW small-world networks and multi-layer forward neural networks, learns from others’ strong points to offset one’s weakness, and considerably improves precision, velocity, and anti-interference ability.

II. SYSTEM FLOW DESCRIPTION

A schematic diagram of the plant is shown in Fig. 1. The crude sewage (CS) from different pumping stations is collected and screened for floating debris and removal of grit is carried out by the grit collector and grit elevators. Primary settlement tanks (PST) are utilized to settle 65–75% of the solids. Settled solids are scrapped down in the hoppers of the PST with the help of mechanical drive scrapers. These settled solids are removed by the Hydro Valves which open in the Consolidation Sludge Tank. Aerobic bacteria are activated by aeration and mixing with activated sludge to reduce the volume of mixed liquor. Primary treated effluent is mixed with the returned activated sludge from the secondary settlement tank and uniformly distributed in channels for aeration with the help of mechanically driven aerators. Mixed liquor out of the aeration tank is made to settle in the secondary settlement tanks. In the post-treatment, the secondary treated effluent is pre-chlorinated and lifted by screw pumps for uniform distribution to sand filters. The resulting stream, designated as final effluent (FE), flows down into the wet well.

![Schematic diagram of the wastewater treatment system](image)

Figure 1. Schematic diagram of the wastewater treatment system

III. CONSTRUCTION OF THE MODEL OF EFFLUENT QUALITY PREDICTION

A. The Input and Output Variables of Model

The biological oxygen demand (BOD) detected at the position E may be affected by the variable value at A, B, C, D, E some hours ago (see Fig.1). There are certain key parameters which can be used to assess the plant performance in a wastewater treatment plant. These parameters could include BOD, suspended solid and chemical oxygen demand.

B. Decision of Hidden Layer Structure

A multi-layer neural network is composed of input, hidden and output layers. In neural network, all nodes per layer are full-connected with those in their adjacent layers. Selecting rational hidden structure is the most important problem in model selection. The NW small-world artificial neural networks will have better performance if the number of connections between nodes in distant layers is very small. So, we can determine the number of neurons and layers in hidden layers small-world artificial neural networks using the methods of general neural network.

Table 1 lists all variables and their corresponding meanings which will be used in this paper.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Meaning</th>
<th>Variable</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>the number of input nodes</td>
<td>N</td>
<td>the number of given samples</td>
</tr>
<tr>
<td>O</td>
<td>the number of output nodes</td>
<td>X_j</td>
<td>the nth input sample</td>
</tr>
<tr>
<td>S</td>
<td>the number of hidden layers</td>
<td>Y_j</td>
<td>the nth teacher output</td>
</tr>
<tr>
<td>M_s</td>
<td>the number of nodes for the nth hidden layer</td>
<td>d_j</td>
<td>the nth real output</td>
</tr>
</tbody>
</table>

In Table 1, s = 1, 2, ..., S and k = 1, 2, ..., N.

The neurons’ weight and threshold are unknown in model of networks, the number of unknown quantities in first hidden layer is \((I + 1)M_1\). Similarly, \((M_n + 1)M_{n+1}\) is the number of unknown quantities in the No. n hidden layer. Then, the quantities can be expressed as

\[
L = (I + 1)M_1 + \sum_{k=1}^{N-1}(M_n + 1)M_{n+1} + (M_n + 1)J
= (I + 1)M_1 + \sum_{k=0}^{N-1}(M_n + 1)M_{n+1} + (M_n + 1)J
\]

\[
M_0 = -1. \quad (1)
\]
Neural networks are usually based on supervised learning. Mean square error function is selected as the objective function of algorithm which can be written as
\[
E(Z) = \frac{1}{2K} \sum_{k=1}^{K} \sum_{j=1}^{J} (d_k^j - y_k^j(Z))^2 .
\]  

Where \(y_k^j(Z)\) is the actual output of the network. Back Propagation works by approximating the non-linear relationship between the input and the output by adjusting the weight values internally along with the gradient direction. Ideally, there exists a \(Z\), let \(E(Z) = 0\), then
\[
E(Z) = \frac{1}{2K} \sum_{k=1}^{K} \sum_{j=1}^{J} (d_k^j - y_k^j(Z))^2 = 0 .
\]  

From (4) we can see that the number of equation is \(JK\) and the number of variables is \(S\). According to algebra equation theory we can get
\[
S = K \cdot J .
\]  

According to equation (1) and (5) let \(S = L\), we can get
\[
(I + 1)M_1 + \sum_{n=0}^{N-1} (M_n + 1)M_{n+1} + (M_N + 1) = K \cdot J .
\]  

For example, we can get the formula to determine the number of nodes per hidden layer then \(S = 1, 2, 3\), as follows

a) When \(S = 1\)

From (6), we can get
\[
(I + 1)M_1 + (M_1 + 1)J = K \cdot J .
\]  

With simplifying the formalities,
\[
M_1 = \text{int} \left[ \frac{J(K-1)}{I + J + 1} \right] .
\]  

b) When \(S = 2\)

From (6), we can get
\[
(I + 1)M_1 + (M_1 + 1)M_2 + (M_2 + 1)J = K \cdot J.
\]  

Then
\[
M_1 = \text{int} \left[ \frac{J(K-1) - (J + 1)M_2}{M_2 + I + 1} \right] .
\]  

Or
\[
M_2 = \text{int} \left[ \frac{J(K-1) - (J + 1)M_1}{M_1 + J + 1} \right] .
\]  

c) When \(S = 3\)

From (6), we can get
\[
(I + 1)M_1 + (M_1 + 1)M_2 + (M_2 + 1)M_3 + (M_3 + 1)J = K \cdot J .
\]  

If we let \(M_1, M_2\) as a definite value, then
\[
M_3 = \text{int} \left[ \frac{J(K-1) - (I + 1)M_1 - (M_1 + 1)M_2}{M_2 + J + 1} \right] .
\]  

According to Kolmogorov theory [8]: A continuous function: \(f : [0,1]^I \rightarrow R^J\), can be achieved by 3 layers front neural network, there are \(I\) neurons in the input layer, \(2I + 1\) neurons in the middle layer, \(J\) neurons in the output layer of the network.

So, we can know
\[
M = 2I + 1, \text{ when } S = 1; \]
\[
M_s < 2I + 1, \text{ when } S = 2, 3, \cdots, S;\]

Where \(M_s\) is the number of nodes for the \(s\)th hidden layer.

Then, we can evaluate the maximal number of nodes for all hidden layers
\[
\sum_{n=1}^{N} M_n \leq 2I + 1 .
\]  

In fact, we usually need very few hidden layers to solve the applications. Having the maximum of hidden layers, the hidden structure can be determined.

In order to get the optimal hidden structure, we compare several network construct. As Table II show.

<table>
<thead>
<tr>
<th>Table II. COMPARISON OF SEVERAL NETWORK CONSTRUCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>the first hidden layer</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

From Table II, we can see the optimal structure; it contains 5 neurons in the input layer, a neuron in the output layer, 2 neurons in the first hidden layer, 3 neurons in the second hidden layer and 8 neurons in the third hidden layer. The network structure has high rate of convergence and a low level of error.
IV. GENERATION AND ALGORITHM OF NW MULTILAYER FEEDFORWARD SMALL-WORLD ARTIFICIAL NEURAL NETWORKS

A. Model Generation Process

Based on the topology of the NW small-world networks, the model of NW multilayer feedforward small-world artificial neural networks is proposed. The model generation process is as follows.

1. Initially, neurons are connected feedforward, i.e. each neuron of a given layer connects to all neurons of the subsequent layer.
2. We make a random draw of two nodes which are connected to each other. We don’t cut that “old” link. In order to create a “new” link, we make another random draw of two nodes. If those nodes are already connected to each other, we make further draws until we find two unconnected nodes. Then we create a “new” link between those nodes.
3. In this way we create some connections between nodes in distant layers, i.e. short-cuts and the topology changes gradually (see Fig. 2).

B. The Flow of the Algorithm

The NW small-world artificial neural networks program-training process as follows:

Step 1: Design the structure of neural network and input parameters of the network.
Step 2: Get initial weights W and initial threshold values from randomizing.
Step 3: Input training data matrix X and output matrix d_k.
Step 4: Compute the output vector of each neural units.
   (a) Compute the input and output vector of the No.1 hidden layer
       \[ U_{m_1}^{M_1} = \sum_{i=1}^{I} W_{im_1}^{M_1} x_{ki} \]  (15)
       \[ v_{m_1}^{M_1} = f(\sum_{i=1}^{I} W_{im_1}^{M_1} x_{ki}) ; m_1 = 1,2,\cdots,M \]  (16)
   (b) Compute the output vector of the output layer neurons

   \[ U_{m_s}^{Ms} = \sum_{i=1}^{I} W_{im_s}^{Ms} x_{ki} + \sum_{i=1}^{I} \sum_{m=1}^{M_i} W_{m,m_s}^{Ms} u_{m_i}^{Ms} \]  (17)
   \[ v_{m_s}^{Ms} = f(u_{m_s}^{Ms}); M_s = 1,2,\cdots,M \]  (18)

   (c) Compute the error signal of the Output layer neurons

\[ y_{ko} = f(u_{o}^{M_o}); o = 1,2,\cdots,O \]  (20)
Step 5: Compute the local gradient
(a) Compute the local gradient of the output layer $M_o$
$$\delta_{o_M}^M(n) = \frac{1}{2} \sum_{i=1}^{O} (d_{i_M}(n) - y_{i_M}(n))^2 .$$ (21)
(b) Compute the local gradient of the last one hidden layer
$$\delta_{m_1}^M(n) = v_{m_1}^M(n)(1 - v_{m_1}^M(n)) \sum_{j=1}^{p} \sum_{i=1}^{O} \delta_{o_j}^M(n)w_{m_1 m_j}(n) .$$ (22)
(c) Compute the local gradient of the hidden layer $m_s$
$$\delta_{m_s}^M(n) = v_{m_s}^M(n)(1 - v_{m_s}^M(n)) \sum_{t=1}^{S} \sum_{j=1}^{p} \delta_{m_t}^M(n)w_{m_s m_j}(n) \sum_{i=1}^{O} \delta_{o_i}^M(n)w_{m_s o_i}(n) .$$ (23)
Step 6: Renew W
$$\Delta w_{ji}^S = w_{ji}^S(n-1) + \eta \delta_{j}^S y_{j}^{S-1}(n-1) .$$ (26)
$$\Delta b_{ji}^S = \alpha b_{ji}^S(n-1) + \eta \delta_{j}^S y_{j}^{S-1}(n-1) .$$ (27)
Step 7: Repeat step 3 to step 6 until converge.

C. Prediction Step of Modle
The prediction process consists of nine steps:
1) Initialization, we let the network contain $L$ hidden layers $n_i$ or $p$ neurons in the hidden layers.
2) The data of input will be normalized.
3) Have circuit training for every sample.
4) Determine whether the cycle has finished. If not, return the step c.
5) The total error $E$ should be computed to check whether the results meet accuracy requirements, if $E < \varepsilon$ , then go on to step f, or go to step c.
6) Whether the iterations exceed the largest number of iterations. If yes, go on to step 7, or go to 3.
7) Record the metrics of the training network saved for the prediction of wastewater treatment.
8) Compute the predicted value of Sewage effluent quality.

i) Denormalization, get the real data of prediction.

V. SIMULATIONS
Established prediction model of NW multilayer feedforward small-world artificial neural networks, a comparative test was made on the regular multilayer feedforward networks and NW small-world artificial neural networks ($p = 0.1$) about convergence speed, precision and stability by using MATLAB 7.0 as Simulation Tool.

In the test, we set up a five layers neural network that has three hidden layers, one input layer and one export layer. From table II, the network consists of three hidden layers, and the first, second and third layers have 2, 3 and 8 neurons respectively. We set the rewiring probability $p = 0.1$, run 100 times independently until up to 10000 iterations. During the realization process of algorithm, we let inertia coefficient $\alpha = 0.9$ , learning velocity $\eta = 0.05$ . There are 5 input and one expected output, let $\varepsilon < 0.01$.

![Figure 3 and 4](https://via.placeholder.com/150)

**Figure 3.** The figure of training following of general ANNs

**Figure 4.** The figure of training following of small-world ANNs

124 valid data have been collected and could be cut into two parts: training sample (80) and predictive sample.
From Fig. 3 and 4, we find that the small-world network with \( p=0.1 \) gives faster convergence compared to the regular network.

Fig 5 and 6 the training result of effluent quality BOD5 of the above two networks.

Fig 7 and 8, we can get the error curve of the two networks above.

Fig 9 and 10 show predicting results of effluent quality BOD5 of the above two networks.
From the comparison of the two networks above, this paper summed up the differences of both in convergence rate, accuracy and robustness: Based on the comparison with the forecast result from regular network, it is demonstrated that the convergence speed of the multilayer feedforward small-world artificial neural network is faster with more accurate precision. Actual values of BOD in the training and testing data are compared to predicted values by the BP neural network models and NW multilayer feedforward small-world artificial neural network, to evaluate the models performance. Visual inspection indicates that the NW multilayer feedforward small-world artificial neural network models resulted in a good fit for the measured BOD data.

### Table III.

<table>
<thead>
<tr>
<th>NETWORK</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>regular multilayer feedforward networks</td>
<td>0.1332</td>
</tr>
<tr>
<td>multilayer feedforward small-world neural networks</td>
<td>0.0010</td>
</tr>
</tbody>
</table>

The table III shows, according to MSE, the multilayer feedforward small-world artificial neural network is more superior to that of regular network in training accuracy and prediction accuracy.

The results of prediction indicate that the multilayer feedforward small-world artificial neural network has good predictive effect for the water quality and has high adaptability. The model of small-world artificial neural network runs quickly, is useful to cope with the fluctuation which could happen at any time. So, it’s an effective method for water quality prediction was provided.

### VI CONCLUSIONS

In this paper, a model based on NW multilayer feedforward small-world artificial neural network is developed to predict the effluent concentrations of BOD for a WWTPs. The model is shown to fit the data precisely and to overcome several disadvantages of the conventional BP neural networks, namely: slow convergence, low accuracy and difficulty in finding the global optimum. A series of tests have been conducted based on the samples. It has been shown that the NW multilayer feedforward small-world artificial neural network model provided good estimates for the BOD data sets. After the network is trained, it becomes simple, fast, and precise, with strong self-adaptability and anti-interference ability to dispose data while predicting the wastewater treatment plant performance. The limitation in data, however, should be highlighted. If more data were collected, if the data were less noisy, this would have resulted in an improved predictive capability of the network.

### ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (grant number: 61040012).

### REFERENCES


Ruicheng Zhang, be born in March 1975, Ph.D. associate professor master instructor in control science and engineering, and Ph.D.in control theory and control engineering, University of science and technology Beijing, mainly engaged in rolling automation, production intelligent control and robust control and other research. He was awarded the Hebei University of Technology Excellent Teacher third prize, outstanding graduate design instructor, excellent teacher, excellent script and other honorary titles. In recent years, he guiding students to participate in ”Freescale” Cup National University Smart Car race, won the first prize and second prize. He has published 30 papers in leading journals such as Journal of Vibration and Control, Control Theory and Application, transactions of china electro technical society, Journal of University of Science and Technology Beijing etc, which has been tagged 11 times by SCI and EI, and cited 50 times by others. Dr. Zhang has co-authored 1 book in Chinese. Associate professor Rui-Cheng Zhang currently chaired by the National Natural Science Foundation: Small world artificial neural network model; Ministry of Science and Technology personnel services business operations funded projects: JDC-150 type electronic digital multi-function tape measure ruler; Natural Science Foundation of Hebei Province: ADRC control theory and applied research of vibration control in the rolling mill drive system and other projects.

Xulei Hu, be born in December 1985, Master's degree student of engineering in control theory and control engineering, Hebei United University, mainly engaged in modeling and simulation of complex industry control system.