Predicting performance of grey and neural network in industrial effluent using online monitoring parameters

(Running title: Grey and network prediction of industrial effluent using online parameters)

T. Y. Pai a,*, S.H. Chuang a, H.H. Ho a, L.F. Yu a, H.C. Su a and H.C. Hu a

a Department of Environmental Engineering and Management, Chaoyang University of Technology, Wufeng, Taichung, 41349, Taiwan, R.O.C.

*Corresponding author: Tzu-Yi Pai

Department of Environmental Engineering and Management
Chaoyang University of Technology, Taichung, 41349, Taiwan
Tel: +886-4-23323000-4465, Fax: +886-4-23742365, E-mail: bai@ms6.hinet.net
Abstract

Grey model (GM) and artificial neural network (ANN) were employed to predict suspended solids (SS$_{\text{eff}}$), chemical oxygen demand (COD$_{\text{eff}}$) and pH$_{\text{eff}}$ in the effluent from conventional activated process of an industrial wastewater treatment plant using simple online monitoring parameters (pH in the equalization pond effluent; pH, temperature, and dissolved oxygen in the aeration tank). The results indicated that the minimum mean absolute percentage errors of 20.79 %, 6.09 % and 0.71 % for SS$_{\text{eff}}$, COD$_{\text{eff}}$ and pH$_{\text{eff}}$ could be achieved using different types of GMs. GM only required a small amount of data (at least 4 data) and the prediction results were even better than those of ANN. According to the results, the online monitoring parameters could be applied on the prediction of effluent quality. It also revealed that GM could predict the industrial effluent variation as its effluent data was insufficient.

Keywords: Grey model; artificial neural network; industrial wastewater treatment plant; conventional activated sludge process; biological treatment; industrial park
1. Introduction

The activated sludge process (ASP) which belongs to the biological wastewater treatment process is broadly used for industrial wastewater treatment plant (IWWTP). When adopting ASP in IWWTPs, it will yield some problems due to the complexity in influent quality and quantity. Literatures have shown that many water quality indices have been investigated to valid mechanistic models. So the more the items for wastewater characterization are, the more the reactions in ASP can be understood [1]. In Taiwan, if the effluent comes from the designated sewers of communities or other residential area, only four effluent characteristics, i.e., suspended solids (SS), biochemical oxygen demand (BOD), chemical oxygen demand (COD) and true color, are regulated according to Effluent Standard. Meanwhile, in order to save cost, effluent investigation from IWWTPs is only carried out to meet regulation standard, so their investigation data are few and incomplete. Under this situation, the effluent cannot be predicted appropriately using some numerical models, especially mechanism models. Some soft computation techniques, such as artificial neural network (ANN), in which the mechanism reactions can be ignored are available presently and applied not only in biotechnology [2-4] but also in biological wastewater treatment process. Although ANN can predict the effluent from WWTPs successfully, many data are required for further calculation [5-9]. Due to the fact that only few parameters are selected for regulation, it is necessary to adopt a suitable method to present these limited data appropriately, the grey system theory (GST) is a suitable method.

The GST proposed by Deng [10] can resolve the problem of incomplete information and data [11-12]. GST requires only a small amount of data and the better prediction results can be obtained. There are many analysis methods in GST including grey relational analysis (GRA) and grey model (GM). In our previous study, we used GM to predict the effluent quality from a hospital WWTP [11]. However, influent quality indices including SS and COD were adopted as the input variables in that study. When determining the concentrations of SS and COD, some manual experiments in laboratory must be carried out, so manual determination of SS and COD cannot respond to operation conditions of WWTP immediately. If adopting the online monitoring equipment to
determine SS or COD, cost will increase because the equipment is expensive. Fortunately, there are many simple and cheap online meters such as DO or pH meters equipped in IWWTP. If the relation between these simple online monitoring parameters and effluent quality can be constructed, a better control strategy could be sought.

The objectives of this study are listed as follows. (1) Calculate the grey relational grades (GRGs) between these simple online monitoring parameters and effluent quality to find out the parameters which affect effluent quality significantly. (2) Use GM to establish the effluent characteristics of an IWWTP, and then the effluent quality was predicted. (3) For comparison, ANN was also employed to predict the effluent in this study.

2. Materials and methods

2.1. Treatment process

The conventional ASP was adopted in this IWWTP. The flow rate was 7500 cubic meters per day (CMD). The online monitoring parameters and effluent quality from 3rd of January 2005 to 6th of October 2005 were investigated. They were sampled and investigated every two to three days and their total numbers were 160. Among the total numbers of data, the numbers for training and testing (predicting) were 130 and 30, respectively. The input variables included pH in the equalization pond effluent \( \text{pH}_{eq} \), pH in the aeration tank \( \text{pH}_{ae} \), temperature in the aeration tank \( \text{Temp}_{ae} \), and DO in the aeration tank \( \text{DO}_{ae} \). The output variables included effluent SS \( \text{SS}_{eff} \), effluent COD \( \text{COD}_{eff} \) and effluent pH \( \text{pH}_{eff} \). All analytical methods used in this study were according to Standard Method [13].

2.2. Grey relational analysis

The aim of GRA is to investigate the factors that affect the system based on finding the relationships of both independent data series. By finding the GRA mathematically, the GRG can be used to evaluate the relational level between referential series and each comparative series. The algorithm of GRA is illustrated as follows. In grey relational space, there is a referential series with \( k \) entities described as: \( X_0(k), k = 1, 2, \ldots, n \). In this study, the effluent quality, i.e. \( \text{SS}_{eff}, \text{COD}_{eff} \) and
\( \text{pH}_{\text{eff}} \), was taken as the referential series. When calculating GRGs, \( n \) is equal to 130. There are also \( m \) comparative data series with \( k \) entities described as: \( X_i(k), i = 1, 2, \ldots, m; k = 1, 2, \ldots, n \). Since there were 4 different online monitoring parameters (\( \text{pH}_{\text{eq}}, \text{pH}_{\text{ae}}, \text{Temp}_{\text{ae}}, \) and \( \text{DO}_{\text{ae}} \)) in this study, \( m = 4 \) when calculating GRGs. That is

\[
\begin{align*}
X_0 &= (x_0(1), x_0(2), \ldots, x_0(n)) \\
X_i &= (x_i(1), x_i(2), \ldots, x_i(n)) \\
X_2 &= (x_2(1), x_2(2), \ldots, x_2(n)) \\
& \vdots \\
X_m &= (x_m(1), x_m(2), \ldots, x_m(n)) 
\end{align*}
\]

The grey relational coefficients (GRCs) between the comparative series \( X_i \) and the referential series \( X_0 \) at the \( k \)-th entity is defined as:

\[
\gamma(X_0, X_i) = \frac{\Delta \min + \varsigma \cdot \Delta \max}{\Delta_{0i}(k) + \varsigma \cdot \Delta \max},
\]

where \( \Delta_{0i}(k) \) is the absolute value of difference between \( X_0 \) and \( X_i \) at the \( k \)-th entity, that is,

\[
\Delta_{0i}(k) = \max_{\forall i} \min_{\forall k} \left| x_0(k) - x_i(k) \right|,
\]

\[
\Delta \min = \min_{\forall i} \min_{\forall k} \Delta_{0i}(k) = \min_{\forall i} \min_{\forall k} \left| x_0(k) - x_i(k) \right|,
\]

\[
\Delta \max = \max_{\forall i} \max_{\forall k} \left| x_0(k) - x_i(k) \right|.
\]

\( \varsigma \in [0, 1] \) is the distinguishing coefficient to control the resolution between \( \Delta \max \) and \( \Delta \min \), typically taken as 0.5. The GRG for series \( X_i \) is given as:

\[
\text{GRG} = \frac{1}{n} \sum_{k=1}^{n} \gamma(X_0, X_i) = \frac{1}{n} \sum_{k=1}^{n} \frac{\Delta \min + \varsigma \cdot \Delta \max}{\Delta_{0i}(k) + \varsigma \cdot \Delta \max},
\]

2.3. Grey modeling process

When information is lacking, using few (at least 4) system information, one can create a GM to describe the behavior of the few outputs. By means of accumulated generation operation (AGO), the disorderly data will behave exponentially such that a first-order differential equation can be used to characterize them. Solving the differential equation will yield a time response solution for prediction. Through inverse accumulated generating operation (IAGO), the forecast can be transformed back to the sequence of original series. A grey modeling process is described as follows.
Assume that the original series of data with $n$ samples is expressed as:

$$X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)),$$

where the superscription $(0)$ of $X^{(0)}$ represents the original series. Let $X^{(1)}$ be the first-order AGO of $X^{(0)}$, whose elements are generated from $X^{(0)}$:

$$X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)), \quad \text{where} \quad x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), \text{for } k = 1, 2, \ldots, n.$$ 

Further operation of AGO can be conducted to reach the $r$-order AGO series, $X^{(r)}$:

$$X^{(r)} = (x^{(r)}(1), x^{(r)}(2), \ldots, x^{(r)}(n)), \quad \text{where} \quad x^{(r)}(k) = \sum_{i=1}^{k} x^{(r-1)}(i), \text{for } k = 1, 2, \ldots, n.$$ 

The IAGO is the inverse operation of AGO. It transforms the AGO-operational series back to the one with a lower order. The operation of IAGO for the first-order series is defined as follows:

$$x^{(0)}(k) = x^{(1)}(k) \quad \text{and} \quad x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1) \text{ for } k = 2, 3, \ldots, n.$$ 

After extending this representation to the IAGO of $r$-order series, we have

$$x^{(r-1)}(k) = x^{(r)}(k) - x^{(r)}(k-1) \text{ for } k = 2, 3, \ldots, n.$$ 

The tendency of AGO can be approximated by an exponential function. Its dynamic behavior is like a form of differential equation. The grey model GM $(h, N)$ thus adopts an $n$-order differential equation to fit the AGO-operational series. The parameters $h$ and $N$ in GM $(h, N)$ denotes the order and the number of variables concerned in the differential equation, respectively. The GM $(h, N)$ can be generally expressed as

$$\sum_{j=0}^{b} a_i \frac{d^i x^{(i)}(k)}{dt^i} = \sum_{j=2}^{b} b_j x^{(j)}(k)$$

where the parameter $a$ is the developing coefficient and $b$ is the grey input. In this study, four different types of GM were adopted, i.e. GM $(1, N)$, GM $(1, 1)$ and rolling GM $(1, 1)$ (RGM $(1, 1)$).

**GM $(1, N)$**. According to the definition of GM $(h, N)$, GM $(1, N)$ is that the order in grey differential equation is equal to 1 and defined as follows:

$$x^{(0)}(k) + a2x^{(1)}(k) = \sum_{j=2}^{N} b_j x^{(j)}(k) = b_2 x^{(1)}(k) + b_3 x^{(1)}(k) + \cdots + b_N x^{(N)}(k)$$

where $z^{(l)}(k) = 0.5x^{(1)}(k-1) + 0.5x^{(1)}(k) \quad k = 2, 3, 4, \ldots, n$. Expanding Equation (5), we have
\[
\begin{align*}
    x_{i}^{(0)}(2)+az_{i}^{(1)}(2) &= b_{2}x_{2}^{(0)}(2)+ \cdots + b_{N}x_{N}^{(0)}(2) \\
    x_{i}^{(0)}(3)+az_{i}^{(1)}(3) &= b_{2}x_{2}^{(0)}(3)+ \cdots + b_{N}x_{N}^{(0)}(3) \\
    \vdots \\
    x_{i}^{(0)}(n)+az_{i}^{(1)}(n) &= b_{2}x_{2}^{(0)}(n)+ \cdots + b_{N}x_{N}^{(0)}(n)
\end{align*}
\]  

(6)

Transforming Equation (6) into matrix form, we have

\[
\begin{bmatrix}
    x_{1}^{(0)}(2) \\
    x_{1}^{(0)}(3) \\
    \vdots \\
    x_{1}^{(0)}(n)
\end{bmatrix}
= 
\begin{bmatrix}
    -z_{1}^{(0)}(2) & x_{2}^{(0)}(2) & \cdots & x_{N}^{(0)}(2) \\
    -z_{1}^{(0)}(3) & x_{2}^{(0)}(3) & \cdots & x_{N}^{(0)}(3) \\
    \vdots & \vdots & \ddots & \vdots \\
    -z_{1}^{(0)}(n) & x_{2}^{(0)}(n) & \cdots & x_{N}^{(0)}(n)
\end{bmatrix}
\begin{bmatrix}
    a \\
    b_{2} \\
    \vdots \\
    b_{N}
\end{bmatrix}
\]  

(7)

Then the coefficients can be estimated by solving matrix, \( \hat{\theta} = (B^{T}B)^{-1}B^{T}Y \),

\[
\begin{align*}
    \hat{\theta} &= 
    \begin{bmatrix}
        a \\
        b_{2} \\
        \vdots \\
        b_{N}
    \end{bmatrix} \\
    Y &= 
    \begin{bmatrix}
        x_{1}^{(0)}(2) \\
        x_{1}^{(0)}(3) \\
        \vdots \\
        x_{1}^{(0)}(n)
    \end{bmatrix} \\
    B &= 
    \begin{bmatrix}
        -z_{1}^{(0)}(2) & x_{2}^{(0)}(2) & \cdots & x_{N}^{(0)}(2) \\
        -z_{1}^{(0)}(3) & x_{2}^{(0)}(3) & \cdots & x_{N}^{(0)}(3) \\
        \vdots & \vdots & \ddots & \vdots \\
        -z_{1}^{(0)}(n) & x_{2}^{(0)}(n) & \cdots & x_{N}^{(0)}(n)
    \end{bmatrix}
\end{align*}
\]

(8)

The \( \hat{\theta} \) values represent the weight of comparative series to the referential series. Additionally, the GM (1, N) model could be used for prediction and described as:

\[
\hat{x}_{i}^{(0)}(k) = \sum_{j=2}^{N} b_{j}x_{j}^{(1)}(k) - az_{i}^{(1)}(k)
\]  

GM (1, 1). If the numbers of comparative series were reduced further, the model was GM (1, 1). All time series values of one specific effluent index were used to establish GM (1, 1). Then the constructed GM (1, 1) was used for prediction.

RGM (1, 1). In GM (1, 1), all time series values of one specific index were used to establish GM (1, 1). While in RGM (1, 1), traditionally the time series data of specific effluent water quality index
used to construct model were the 4 data before the point which was considered to be predicted. That is, the model had to be constructed every time step and only 4 data were used for model construction. In both GM (1, 1) and RGM (1, 1), the online monitoring parameter was ignored.

2.4. ANN

The ANN modeling approach simulated the operation features of human nervous system. Many simple computational elements called artificial neurons that are connected by variable weights are used. A typical neural network model consists of three independent layers: input, hidden, and output layers. Each layer is comprised of several operating neurons. Input neurons receive the values of input variables that are fed to the network and store the scaled input values, while the calculated results in output layer are assigned by the output neurons. The hidden layer performs an interface to fully interconnect input and output layers. Each neuron is connected to every neuron in adjacent layers before being introduced as input to the neuron in the next layer by a connection weight, which determines the strength of the relationship between two connected neurons. Each neuron sums all of the inputs that it receives and the sum is converted to an output value based on a predefined activation, or transfer, function. In this study, the back-propagation algorithm was used for teaching ANN which consisted of three independent layers: input, hidden, and output layers. To compare with GM, 2 parameters with higher GRGs (ANN2-1), 3 parameters with higher GRGs (ANN3-1) and all 4 parameters (ANN4-1) were taken as the input layer variables, respectively. Yuan and Vanrolleghem [14] proposed that a rolling training-prediction could be applied in ANN. For comparison, the structure and process that resembled RGM (1, 1) was also adopted and called rolling ANN (RANN) in this work. In RANN, the time series data of specific effluent water quality index used to construct model were the 4 data before the point which was considered to be predicted. That is, the model had to be constructed every time step and only 4 data were used for model construction. The online monitoring parameter was ignored too. Meanwhile each effluent quality was the single output layer variable. After testing, it was found that the optimal numbers of operating neurons in hidden layer was 10. Their calculation was carried out using MATLAB.
2.5. Error analysis

In order to evaluate the prediction accuracy of GM and ANN, the mean absolute percentage error (MAPE) was employed and described

\[
\text{MAPE} = \frac{1}{n} \sum_{k=1}^{n} \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%
\]

where \( x^{(0)}(k) \) is the investigation value, \( \hat{x}^{(0)}(k) \) is the prediction value.

3. Results and discussion

The GRGs between the effluent quality (\( \text{SS}_{\text{eff}} \), \( \text{COD}_{\text{eff}} \) and \( \text{pH}_{\text{eff}} \)) and 4 different online monitoring parameters (\( \text{pH}_{\text{eq}} \), \( \text{pH}_{\text{ae}} \), \( \text{Temp}_{\text{ae}} \), and \( \text{DO}_{\text{ae}} \)) were calculated as follows. The GRGs of \( \text{SS}_{\text{eff}} \) were in the order: \( \text{pH}_{\text{ae}} (0.8814) > \text{pH}_{\text{eq}} (0.8807) > \text{DO}_{\text{ae}} (0.8782) > \text{Temp}_{\text{ae}} (0.6135) \). Those of \( \text{COD}_{\text{eff}} \) were in the order: \( \text{Temp}_{\text{ae}} (0.8852) > \text{pH}_{\text{ae}} (0.7011) > \text{pH}_{\text{eq}} (0.6999) > \text{DO}_{\text{ae}} (0.6957) \). Those of \( \text{pH}_{\text{eff}} \) were in the order: \( \text{pH}_{\text{ae}} (0.9929) > \text{pH}_{\text{eq}} (0.9882) > \text{DO}_{\text{ae}} (0.9669) > \text{Temp}_{\text{ae}} (0.5559) \). Based on the results of GRA, the selected input variables in different models were shown in Table 1. Fig. 1 depicts their structures.

Fig. 2 (a), (b), (c), (d), and (e) depict the prediction results using GM1N2-1, GM1N3-1, GM1N4-1, GM (1, 1), and RGM (1, 1), respectively. Fig. 3 (a), (b), (c) and (d) show the prediction results using ANN2-1, ANN3-1, ANN4-1, and RANN, respectively. Table 2 shows their MAPE.

When predicting \( \text{SS}_{\text{eff}} \) with GMs, the MAPE of 20.79 % was the lowest when using both GM1N2-1 and RGM (1, 1). This value was lower than those of ANN2-1, ANN3-1, ANN4-1, and RANN by 4.99 %, 1.48 %, 4.74 % and 3.10 %, respectively. When predicting \( \text{COD}_{\text{eff}} \) with GMs, the MAPE of 6.09 % was the lowest when using RGM (1, 1). It was lower than those of ANN2-1, ANN3-1, ANN4-1 and RANN by 5.80 %, 6.70 %, 5.41 % and 0.81 %, respectively. When predicting \( \text{pH}_{\text{eff}} \) with GMs, the MAPE of 0.71 % was the lowest when using RGM (1, 1). This value was lower than those of ANN2-1, ANN3-1, ANN4-1 and RANN by 0.61 %, 0.55 %, 0.50 % and 0.80 %, respectively.

Additional testing was made for further comparison. The 9 month data were split into different
training-testing sets: 1-3, 5-8th months as training data, while 4th and 9th months as testing data.

The results indicated that RGM (1, 1) still performed better (Table 2). The reason could be explained as follows. RGM (1, 1) only used 4 data before the point which was considered to be predicted. But when training (or constructing) ANN, all data should be used even in RANN. More data would result in higher prediction error. Usually, training error was lower than prediction error.

In Table 2, it was found that training error was higher than prediction error in some cases. It could be explained based on standard deviation which measured the spread of the data about the mean value. When training (130 data), the values of standard deviation were 4.00, 6.28 and 0.18 for SS$_{eff}$, COD$_{eff}$ and pH$_{eff}$, respectively. When testing (30 data), those of SS$_{eff}$ COD$_{eff}$ and pH$_{eff}$ were 1.92, 3.19 and 0.10, respectively. The variation of standard deviation resulted in the phenomenon in which the training error was higher than prediction error.

4. Conclusions

The simulation results can be drawn as follows. The minimum MAPEs of 20.79 %, 6.09 % and 0.71 % for SS$_{eff}$, COD$_{eff}$ and pH$_{eff}$ could be achieved using GMs. A good fitness could be achieved using ANN too, but they required a large quantity of data for constructing model. Contrarily, GM only required a small amount of data (at least 4 data) and the prediction results were even better than those of ANN. The results also indicated that RGM (1, 1) still performed better than ANN even different training-testing sets were used. According to the results, the online monitoring parameters could be applied on prediction of effluent. It also revealed that GM could predict the industrial effluent variation as its effluent data was insufficient.

Acknowledgements

The authors are grateful to the National Science Council of Taiwan, R.O.C. for financial support under the grant number NSC 94-2211-E-324-007.
11

References


Chem Eng 2007;31(10),1272-81.


<table>
<thead>
<tr>
<th></th>
<th>GM</th>
<th>ANN</th>
<th>Input variables</th>
<th>ANN</th>
<th>Input variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SS$_{eff}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM1N2-1</td>
<td>pH$<em>{ac}$, pH$</em>{eq}$</td>
<td>ANN2-1</td>
<td>pH$<em>{ac}$, pH$</em>{eq}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM1N3-1</td>
<td>pH$<em>{ac}$, pH$</em>{eq}$, DO$_{ae}$</td>
<td>ANN3-1</td>
<td>pH$<em>{ac}$, pH$</em>{eq}$, DO$_{ae}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM1N4-1</td>
<td>pH$<em>{ac}$, pH$</em>{eq}$, DO$<em>{ae}$, Temp$</em>{ae}$</td>
<td>ANN4-1</td>
<td>pH$<em>{ac}$, pH$</em>{eq}$, DO$<em>{ae}$, Temp$</em>{ae}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM (1, 1)</td>
<td>all SS$_{eff}$ for training</td>
<td>-</td>
<td>-</td>
<td>RANN</td>
<td>previous 4 SS$_{eff}$ before predicted point</td>
</tr>
<tr>
<td>RGM (1,1)</td>
<td>previous 4 SS$_{eff}$ before predicted point</td>
<td>RANN</td>
<td>previous 4 SS$_{eff}$ before predicted point</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>COD$_{eff}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM1N2-1</td>
<td>Temp$<em>{ac}$, pH$</em>{ac}$</td>
<td>ANN2-1</td>
<td>Temp$<em>{ac}$, pH$</em>{ac}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM1N3-1</td>
<td>Temp$<em>{ac}$, pH$</em>{ac}$, pH$_{eq}$</td>
<td>ANN3-1</td>
<td>Temp$<em>{ac}$, pH$</em>{ac}$, pH$_{eq}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM1N4-1</td>
<td>Temp$<em>{ac}$, pH$</em>{ac}$, pH$<em>{eq}$, DO$</em>{ae}$</td>
<td>ANN4-1</td>
<td>Temp$<em>{ac}$, pH$</em>{ac}$, pH$<em>{eq}$, DO$</em>{ae}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM (1, 1)</td>
<td>all COD$_{eff}$ for training</td>
<td>-</td>
<td>-</td>
<td>RANN</td>
<td>previous 4 COD$_{eff}$ before predicted point</td>
</tr>
<tr>
<td>RGM (1,1)</td>
<td>previous 4 COD$_{eff}$ before predicted point</td>
<td>RANN</td>
<td>previous 4 COD$_{eff}$ before predicted point</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>pH$_{eff}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM1N2-1</td>
<td>pH$<em>{ac}$, pH$</em>{eq}$</td>
<td>ANN2-1</td>
<td>pH$<em>{ac}$, pH$</em>{eq}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM1N3-1</td>
<td>pH$<em>{ac}$, pH$</em>{eq}$, DO$_{ae}$</td>
<td>ANN3-1</td>
<td>pH$<em>{ac}$, pH$</em>{eq}$, DO$_{ae}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM1N4-1</td>
<td>pH$<em>{ac}$, pH$</em>{eq}$, DO$<em>{ae}$, Temp$</em>{ae}$</td>
<td>ANN4-1</td>
<td>pH$<em>{ac}$, pH$</em>{eq}$, DO$<em>{ae}$, Temp$</em>{ae}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM (1, 1)</td>
<td>all pH$_{eff}$ for training</td>
<td>-</td>
<td>-</td>
<td>RANN</td>
<td>previous 4 pH$_{eff}$ before predicted point</td>
</tr>
<tr>
<td>RGM (1,1)</td>
<td>previous 4 pH$_{eff}$ before predicted point</td>
<td>RANN</td>
<td>previous 4 pH$_{eff}$ before predicted point</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2 MAPEs between the predicted and investigated values using different GM and ANN

<table>
<thead>
<tr>
<th></th>
<th>SS&lt;sub&gt;eff&lt;/sub&gt;</th>
<th>COD&lt;sub&gt;eff&lt;/sub&gt;</th>
<th>pH&lt;sub&gt;eff&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Construction</td>
<td>Prediction</td>
<td>Construction</td>
</tr>
<tr>
<td>GM1N2-1</td>
<td>62.95 %</td>
<td>20.79 %</td>
<td>24.78 %</td>
</tr>
<tr>
<td>GM1N3-1</td>
<td>61.14 %</td>
<td>22.12 %</td>
<td>20.88 %</td>
</tr>
<tr>
<td>GM1N4-1</td>
<td>58.47 %</td>
<td>32.88 %</td>
<td>17.31 %</td>
</tr>
<tr>
<td>GM (1, 1)</td>
<td>59.34 %</td>
<td>36.98 %</td>
<td>15.48 %</td>
</tr>
<tr>
<td>RGM (1, 1)</td>
<td>20.79 %</td>
<td>6.09 %</td>
<td>0.71 %</td>
</tr>
<tr>
<td>RGM (1, 1)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>21.46 %</td>
<td>6.03 %</td>
<td>1.79 %</td>
</tr>
<tr>
<td>ANN2-1</td>
<td>27.34 %</td>
<td>25.78 %</td>
<td>15.78 %</td>
</tr>
<tr>
<td>ANN3-1</td>
<td>23.78 %</td>
<td>22.27 %</td>
<td>15.47 %</td>
</tr>
<tr>
<td>ANN4-1</td>
<td>27.36 %</td>
<td>25.64 %</td>
<td>15.01 %</td>
</tr>
<tr>
<td>RANN</td>
<td>26.18 %</td>
<td>23.89 %</td>
<td>9.17 %</td>
</tr>
<tr>
<td>RANN&lt;sup&gt;a&lt;/sup&gt;</td>
<td>26.00 %</td>
<td>25.40 %</td>
<td>15.09 %</td>
</tr>
</tbody>
</table>

<sup>a</sup>: The 9 month data were split into different training-testing sets.
Fig. 1. The structure diagram of GM and ANN.
Fig. 2. Prediction results using different GM. (a) GM1N2-1 (b) GM1N3-1 (c) GM1N4-1 (d) GM (1, 1) (e) RGM (1, 1).
Fig. 3. Prediction results using different ANN. (a) ANN2-1 (b) ANN3-1 (c) ANN4-1 (d) RANN.