Prediction of biochemical oxygen demand at the upstream catchment of a reservoir using adaptive neuro fuzzy inference system

Yung-Chia Chiu, Chih-Wei Chiang and Tsung-Yu Lee

ABSTRACT

The aim of this study is to examine the potential of adaptive neuro fuzzy inference system (ANFIS) to estimate biochemical oxygen demand (BOD). To illustrate the applicability of ANFIS method, the upstream catchment of Feitsui Reservoir in Taiwan is chosen as the case study area. The appropriate input variables used to develop the ANFIS models are determined based on the t-test. The results obtained by ANFIS are compared with those by multiple linear regression (MLR) and artificial neural networks (ANNs). Simulated results show that the identified ANFIS model is superior to the traditional MLR and nonlinear ANNs models in terms of the performance evaluated by the Pearson coefficient of correlation, the root mean square error, the mean absolute percentage, and the mean absolute error. These results indicate that ANFIS models are more suitable than ANNs or MLR models to predict the nonlinear relationship within the variables caused by the complexity of aquatic systems and to produce the best fit of the measured BOD concentrations. ANFIS can be seen as a powerful predictive alternative to traditional water quality modeling techniques and extended to other areas to improve the understanding of river pollution trends.

Key words | adaptive neuro fuzzy inference system, artificial neural networks, biochemical oxygen demand, Feitsui Reservoir, multiple linear regression, water quality

INTRODUCTION

Hydrodynamic and biochemical processes constitute an integral part of the natural aquatic environment and determine the water quality in an aquatic ecosystem. Prediction of river water quality allows adequate measure to be taken to keep pollutants within permissible limits (Ahmad et al. 2003). The determination of water quality is traditionally based on the consideration of physical, chemical, and biological characteristics and the water usage follows the context of the national or international standard. According to these standards, in the case of water quality, monitoring involves analyzing a large number of variables such as pH, temperature (T), turbidity (TU), suspended solids (SS), dissolved oxygen (DO), electrical conductivity (EC), chemical oxygen demand (COD, derived using potassium dichromate, K2Cr2O7, as the oxidizing agent), total dissolved solid (TDS), ammonia nitrogen (AN), biochemical oxygen demand (BOD), and so on.

In general, the water quality parameters, such as nutrients (nitrogen, phosphorus) or sediments, are considered as the primary targets needed for development of management strategies for the protection of water resources. However, BOD is widely accepted as a surrogate parameter to represent the degree of organic pollution in an aquatic system and is listed as a conventional pollutant. The discharge of wastes with high levels of BOD can cause severe depletion of DO and endanger species living in receiving water bodies. Consequently, it is necessary to monitor BOD concentration in real-time and to control the status of water quality. Performing the test for BOD requires significant time and commitment for preparation and analysis. The entire process requires 5 days, with data collection and evaluation occurring on the last day; therefore, it is referred to as a BOD5 measurement (Delzer & McKenzie 1999). It is well accepted that obtaining the BOD5 is more difficult than other common water quality parameters, such as pH, T, TU, and EC. An accurate well-timed measurement of BOD5 is essential to the successful monitoring and
controlling of water supply systems. Due to the limitations of hardware sensors for online measurements of BOD$_5$, one of the growing developments for the prediction of water quality is mathematical modeling (Roide & Adrian 2007). However, an inherent problem in mathematical modeling is the requirement of model manipulation. Little knowledge is available to generate different input data for modeling process, particularly for BOD$_5$ controlled by many complicated processes. There is no satisfactory explanatory model that relates the prediction variable, i.e. BOD in this study, to other explanatory variables (Lee & Park 1999).

Accordingly, the development and current progress in the integration of various artificial intelligence (AI) techniques, i.e. knowledge-based system, genetic algorithm, artificial neural network (ANN), and fuzzy inference system (FIS), in water quality modeling have been gradually considered and studied by many researchers. Among various AIs, ANNs have been found to be well suited to model a system for which a complete description of the complex input and output relationship remains uncertain. The concept of ANNs was first introduced by McCulloch & Pitta (1943), and the major applications have arisen only since the development of the back-propagation method of training by Rumelhart et al. (1986). ANNs have been widely applied in engineering fields, and successfully used in hydrological processes (Hsu et al. 1995; Srinivasulu & Jain 2006), precipitation (Aksoy & Dahamseh 2009), streamflow (Kisi 2007; Modares 2009; Gao et al. 2010), water resources management and reservoir operations (Wen & Lee 1998; Aguilera et al. 2001), sediment yield and concentration (Nagy et al. 2002; Raghuvanshi et al. 2006), and water quality (Lek et al. 1996; Maier & Dandy 1996; Lek et al. 1999; ASCE Task Committee 2000; Suen et al. 2003; Oliveira-Esquerrre et al. 2004; Sengorur et al. 2006; Dogan et al. 2009; Singh et al. 2009; Zagoot et al. 2009; Ay & Kisi 2012; Wen et al. 2013).

Recently, adaptive neuro fuzzy inference system (ANFIS), which consists of the ANNs and fuzzy logic methods, has been accepted as an efficient alternative tool for modeling of complex hydrologic systems and gotten attention in hydrological modeling and reservoir operation. Fuzzy logic and fuzzy theory were first developed by Zadeh (1965) and applied to water resources forecasting. ANFIS uses the learning ability of the ANNs to define the input-output relationship and construct the fuzzy rules by determining the input structure (Jang 1993). The system results are obtained by the thinking and reasoning capability of the fuzzy logic (Firat & Gungor 2008). ANFIS has been used for prediction of streamflow (Nayak et al. 2004; Firat & Gungor 2008; Kumar et al. 2013), reservoir inflow or water level (Chang & Chang 2006; Lohani et al. 2012), flood forecasting (Chen et al. 2006; Mukerji et al. 2009; Seckin 2011), drought forecasting (Bacanli et al. 2009), sediment yield (Firat & Gungor 2010), and water quality (Civelekoglu et al. 2009; Shekofteh et al. 2013; Zounemat-Kermani & Scholz 2014). Lohani et al. (2012) used the ANFIS method to predict reservoir inflow of the Sutlej River at Bhakra Dam site and indicated that ANFIS can provide more accurate inflow forecasting than ANNs. Chen & Liu (2014) developed an ANFIS model to estimate the DO concentrations in a reservoir and concluded that the ANFIS was superior to the back propagation neural network (BPNN). From the previous discussion it is seen that many researchers have studied the applicability of ANNs and ANFIS models in solving various hydrologic problems and compared their performance under different statistical criteria. However, a comprehensive study to investigate the capability of ANFIS method for long-term monthly water quality forecasting of BOD$_5$ concentrations, and to compare its performance to ANNs and multiple linear regression (MLR) models are still missing.

Therefore, in this study, the main purpose is to present the ANFIS model to forecast BOD$_5$. To verify the application of this approach, the upstream catchment of Feitsui Reservoir located at the northeast part of Taiwan is chosen as the case study area.

**METHODOLOGY**

In this section, the basic concepts and modeling process of proposed ANFIS models are briefly reviewed. The statistical criteria used to evaluate the model performance between the simulated and measured data are also described.

**Adaptive neuro fuzzy inference system**

FIS is a rule-based system consisting of three conceptual components. These are: (1) a rule-base, containing fuzzy if-then rules, (2) a data-base, defining the membership functions (MF), and (3) an inference system, combining the fuzzy rules and producing the system results (Takagi & Sugeno 1985). There are two types of widely used FIS, Takagi-Sugeno FIS and Mamdani FIS (Jang et al. 1997). In this study, Takagi-Sugeno FIS is used for BOD$_5$ prediction. Figure 1 presents a five-layer structure of ANFIS with two input variables. For the first-order Takagi-Sugeno FIS, two typical rules can be expressed as:

---

**Figure 1:** Five-layer structure of ANFIS with two input variables.
Rule 1: If \( x_1 \) is \( A_1 \) and \( x_2 \) is \( B_1 \), then \( f_1 = p_1 x_1 + q_1 x_2 + r_1 \)

Rule 2: If \( x_1 \) is \( A_2 \) and \( x_2 \) is \( B_2 \), then \( f_1 = p_2 x_1 + q_2 x_2 + r_2 \)

where \( x_1 \) and \( x_2 \) are the inputs, \( A_i \) and \( B_i \) are the linguistic labels, \( f_i \) is the output function, \( p_i, q_i, \) and \( r_i \) are the consequent parameters (output function parameters).

Layer 1: each node in this layer generates membership grades of the crisp inputs and each node’s output \( O_i^1 \) is calculated by:

\[
Q_i^1 = \mu_{A_i}(x_1) \quad \text{for} \quad i = 1, 2
\]

\[
Q_i^1 = \mu_{B_i}(x_2) \quad \text{for} \quad i = 3, 4
\]

where \( O_i^1 \) denotes the output of the \( i \)th node, and \( \mu_{A_i} \) and \( \mu_{B_i} \) are the MF for \( A_i \) and \( B_i \) linguistic labels, respectively. The MF can be any appropriate function that is continuous and piecewise differentiable such as Gaussian, bell-shaped, trapezoidal-shaped, and triangular-shaped functions. For example, the bell-shaped function can be expressed as follows:

\[
\mu_{A_i}(x) = 1/1 + \left[ \frac{x - c_i}{a_i} \right]^2
\]

(2)

where \( a_i, b_i, \) and \( c_i \) are the premise parameter set and with the change of these parameters the bell-shaped function varies accordingly.

Layer 2: every node in this layer multiplies the incoming signals, and the output represents the firing strengths of a rule, \( Q_i^2 \) or \( w_i \), is computed as:

\[
Q_i^2 = w_i = \mu_{A_i}(x_1) \mu_{B_i}(x_2), \quad i = 1, 2
\]

Layer 3: this layer calculates the ratio of the \( i \)th rule’s firing strength to the sum of all rules’ firing strengths. It normalizes the firing strength and calculated as:

\[
Q_i^3 = \bar{w}_i = \frac{w_i}{\sum_{i=1}^{2} w_i}, \quad i = 1, 2
\]

Layer 4: this layer calculates only the sum of the signals of the third and second layers of the network

\[
Q_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x_1 + q_i x_2 + r_i), \quad i = 1, 2
\]

Layer 5: this layer computes the output, also called output node, by summing all incoming signals from the fourth layer

\[
Q_i^5 = \sum_i \bar{w}_i f_i = \sum_i \bar{w}_i f_i
\]

A hybrid-learning algorithm which combines the gradient descent and least squares methods to identify parameters defined in the adaptive networks has been proposed. The detailed algorithm and mathematical background of these algorithms can be found in Zadeh (1965), Jang et al. (1997) and Nayak et al. (2004).

Performance criteria

To determine the performance of each identified ANFIS, MLR, and ANNs model, four different criteria are used: the Pearson coefficient of correlation (R), the root mean square error (RMSE), the mean absolute percentage error (MAPE), and the mean absolute error (MAE). Each criterion can be computed as:

\[
R = \frac{\sum_{t=1}^{T_0} (\hat{y}_t - \bar{y})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^{T_0} (\hat{y}_t - \bar{y})^2 \sum_{t=1}^{T_0} (y_t - \bar{y})^2}}
\]

(7)

\[
RMSE = \sqrt{\frac{1}{T_0} \sum_{t=1}^{T_0} (\hat{y}_t - y_t)^2}
\]

(8)

\[
MAPE = \frac{1}{T_0} \sum_{t=1}^{T_0} \frac{|y_t - \hat{y}_t|}{y_t} \times 100
\]

(9)

\[
MAE = \frac{1}{T_0} \sum_{t=1}^{T_0} |y_t - \hat{y}_t|
\]

(10)

where \( t \) is the time index, and \( \hat{y}_t \) and \( y_t \) are the simulated and measured values of BOD\(_5\) at time period \( t \), respectively.
\( \bar{y} \) and \( \bar{y} \) represent the averages of \( y_i \) and \( y_r \), respectively. \( T_0 \) denotes the number of measurements. The higher value of \( R \) and smaller values of RMSE, MAPE, and MAE ensure the better performance of model.

**STUDY AREA AND WATER QUALITY DATA**

In this study, the monthly BOD\( _5 \) measured at the upstream catchment of Feitsui Reservoir in Taiwan are used to develop the ANFIS model. Feitsui Reservoir, located at the downstream of Peishih Creek in the northern part of Taiwan (Figure 2), was built in 1987. It is a water-supply reservoir for the Taipei metropolitan area with a population of over five million. Its surface area is 10.24 km\(^2\) and the mean depth is 39.68 m with a maximum of 113.4 m near the main dam. The climate at the location of reservoir is a subtropical oceanic climate with annual temperature between 18 and 23 °C. The mean annual precipitation and inflow according to historical data from 1989 to 2012 are 3,784 and 3,214 mm, respectively. The reservoir is mainly recharged by the Peishih Creek which has two main tributary creeks, Jin-gua-liu and Dai-yu-ku Creeks (Figure 2). Due to the excessive pollutant loads from municipal, industrial wastewater, and agricultural run-off at the upstream (Figure 2), a decline of water quality is observed in the reservoir (Taipei Water Management Committee 2003).

According to the Carlson trophic state index for water quality of the reservoir, the water quality condition in the Feitsui Reservoir is in the range of mesotrophic and eutrophic (Kuo et al. 2005).

The water quality data used in this study were collected at eight different sampling stations located at the upstream catchment of Feitsui Reservoir over a period of 1987–2012, and the sampling frequency is once per month. Although more than 30 water quality parameters are available, only eight parameters are selected due to their consistency and continuity in measurement at all selected sampling stations. The selected water quality parameters included pH, T, TU, AN, DO, SS, COD, and BOD\( _5 \). Three sampling stations, Ping-lin, Dai-yu-ku Creek, and Jin-gua-liao Creek (Figure 2), are managed by Taipei Feitsui Reservoir Administration and the total number of measured data sets are 936 (312 for each station) from 1/1987–12/2012. Each data set consists of six different water quality parameters, pH, T, TU, AN, DO, and BOD\( _5 \). Five stations, Kuo-lai, Bi-hu, Shui-yuan Bridge, Da-lin Bridge, and Huang-ju-pi-liao (Figure 2), are managed by Taipei Water Management Committee and the total number of measured data sets are 1345 (269 for each station) from 1/1987–5/2009. Each data set consists of eight different water quality parameters, pH, T, TU, AN, DO, SS, COD, and BOD\( _5 \). The water quality parameters of SS and COD measured at Ping-lin, Dai-yu-ku Creek, and Jin-gua-liao Creek stations

![Figure 2](https://i.imgur.com/3Q5Q5Q5.png)

*Figure 2 | Location map of study area and water quality sampling stations.*
were not included in the model development because of their discontinuous monitoring (more than 35% of data were missing). For the model development, the water quality parameters of pH, T, TU, AN, DO, SS, and COD are considered as the explanatory (input) variables \( x_1 - x_7 \) and BOD\(_5\) is seen as the response (output) variables \( y \).

Basic descriptive statistics of each measured water quality parameter at all sampling stations are summarized in Table 1. From Table 1, the explanatory variables show high variability in the standard deviation (SD) and coefficient of variation (CV). Such variability may be attributed to the natural and non-point sources, i.e. topsoil, leaves and woody debris, animal manure, and fertilizer, distributed in the large geographical area of watershed, all of which are very difficult to be quantified. The response variable of BOD\(_5\) also shows high variability in the SD and CV and has a maximum value of 4.3 mg/L at Ping-lin station. The mean value for each station ranges from 0.55 to 0.92 mg/L (not shown). The correlation coefficients between BOD\(_5\) and the explanatory variables were also calculated and presented in Table 1. The coefficients show that the linear correlation between BOD\(_5\) and the explanatory variables is low. The highest two values of 0.317 and 0.202 are between BOD\(_5\) and COD, BOD\(_5\) and T, respectively. This result implies that using only a linear model, i.e. MLR, to establish the underlying relationship between BOD\(_5\) and the other water quality parameters might be insufficient. Besides, BOD\(_5\) is the amount of oxygen required for microbial metabolism of organic compounds in water, and the rate of oxygen consumption depends on T, pH, the presence of certain kinds of microorganisms, the type of organic material, and the enzymes available to indigenous microbial populations (Karube et al. 1977; Jouanneau et al. 2014; Heddam et al. 2016). If the underlying (physical, chemical, and biological) processes of BOD\(_5\) are considered, there is no single variable which can completely represent the variations of BOD\(_5\) in an aquatic ecosystem, as suggested by the low correlations shown in Table 1.

The spatial (site-wise) data sets of measured water quality parameters are statistically compared for any significant difference among the sampling stations. The calculated statistical values of the F-test are higher than the respective critical value at 95% probability level and indicate that there is a significant difference among the stations (not shown). The difference among the stations may also be attributed to the fact that in the upper-course (Kuo-lai and Bi-hu stations, Figure 2), the river water quality is mainly dominated by the variables of geogenic origin, and there are no major pollution sources in the region. In the mid- or lower-course (Jin-gua-liau Creek, Shui-yuan Bridge, and Huang-ju-pi-liao stations, Figure 2), the river receives heavy loads of untreated wastewater from nearby townships and is dominated by the variables of anthropogenic origin.

**RESULTS OF ANFIS MODEL**

In this section, the water quality data at the upstream catchment of a reservoir are used to demonstrate the developments of ANFIS models. For ANFIS identification, the entire data set for each station is divided into two sets, the training and testing sets. Seventy percent of the data are used for the training, and 30% are used for the testing. For the Ping-lin, Dai-yu-ku Creek, and Jin-gua-liau Creek stations, Figure 2, the river receives heavy loads of untreated wastewater from nearby townships and is dominated by the variables of anthropogenic origin.

<table>
<thead>
<tr>
<th>Parameter (unit)</th>
<th>Observed station</th>
<th>Observation period</th>
<th>Max</th>
<th>Min</th>
<th>Mean</th>
<th>SD</th>
<th>CV</th>
<th>Cor_BOD(_5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 ) - pH ((^{-}))</td>
<td>ALL</td>
<td>1987-2012</td>
<td>9.60</td>
<td>6.00</td>
<td>7.51</td>
<td>0.50</td>
<td>6.62</td>
<td>0.037</td>
</tr>
<tr>
<td>( x_2 ) - T ((^\circ)C)</td>
<td>ALL</td>
<td>1987-2012</td>
<td>34.00</td>
<td>7.30</td>
<td>21.35</td>
<td>4.66</td>
<td>21.83</td>
<td>0.202</td>
</tr>
<tr>
<td>( x_3 ) - TU (NTU)</td>
<td>ALL</td>
<td>1987-2012</td>
<td>205.00</td>
<td>0.00</td>
<td>3.28</td>
<td>9.13</td>
<td>278.56</td>
<td>-0.011</td>
</tr>
<tr>
<td>( x_4 ) - AN (mg/l)</td>
<td>ALL</td>
<td>1987-2009</td>
<td>1.56</td>
<td>0.00</td>
<td>0.04</td>
<td>0.05</td>
<td>125.54</td>
<td>0.086</td>
</tr>
<tr>
<td>( x_5 ) - DO (mg/l)</td>
<td>ALL</td>
<td>1987-2009</td>
<td>11.20</td>
<td>1.00</td>
<td>8.67</td>
<td>0.87</td>
<td>10.08</td>
<td>-0.033</td>
</tr>
<tr>
<td>( x_6 ) - SS(^a) (mg/l)</td>
<td>Kuo-lai, Bi-hu, Da-lin Bridge, Shui-yuan Bridge, Huang-ju-pi-liao</td>
<td>1987-2009</td>
<td>263.30</td>
<td>0.20</td>
<td>3.05</td>
<td>8.81</td>
<td>289.30</td>
<td>0.069</td>
</tr>
<tr>
<td>( x_7 ) - COD(^b) (mg/l)</td>
<td>Kuo-lai, Bi-hu, Da-lin Bridge, Shui-yuan Bridge, Huang-ju-pi-liao</td>
<td>1987-2009</td>
<td>18.30</td>
<td>0.10</td>
<td>2.73</td>
<td>1.94</td>
<td>71.13</td>
<td>0.317</td>
</tr>
<tr>
<td>( y ) - BOD(_5) (mg/l)</td>
<td>ALL</td>
<td>1987-2009</td>
<td>4.30</td>
<td>0.00</td>
<td>0.70</td>
<td>0.52</td>
<td>74.01</td>
<td>1.00</td>
</tr>
</tbody>
</table>

\( x_i \) – independent variable; \( y \) – dependent variable; Max – Maximum; Min – Minimum; SD – Standard deviation; CV – Coefficient of variation; Cor_BOD\(_5\) – Correlation with BOD\(_5\).

\(^a\)SS and COD data measured at Ping-lin, Dai-yu-ku Creek, and Jin-gua-liau Creek stations are not included in the model development.
stations, the data from 1/1987–2/2005 (218 data sets for each station) are used for the training, and the data from 3/2005–12/2012 (94 data sets for each station) are used for the testing. For the Kuo-lai, Bi-hu, Shui-yuan Bridge, Da-lin Bridge, and Huang-ju-pi-liao stations, the data from 1/1987–8/2002 (188 data sets for each station) are used for the training, and the data from 9/2002–5/2009 (81 data sets for each station) are used for the testing.

In view of the requirements of the computation algorithm, the raw data of both the input and output variables are normalized to an interval by transformation before training the system. Here, all the water quality data are transformed to the same uniform distribution between −1 and +1. In order to effectively identify the optimal ANFIS architecture, the selection of input variables (nodes), the number of MFs, and the selection of MFs are determined cautiously and thoroughly. The five-layer ANFIS model identified by the hybrid-learning algorithm is developed for each station and the performance of all identified models are evaluated by R, RMSE, MAPE, and MAE.

**ANFIS(I)**

For ANFIS identifications, the selection of input variables, a very important aspect of the ANFIS modeling, is traditionally based on either an ad hoc basis or trial-and-error method. However, these selection processes are very time consuming and unreliable. Therefore, the determination of input variables based on the t-test, typically used to test the significance of an individual coefficient in the MLR model, is applied in this study. The significant variables obtained from the t-test imply their importance to the BOD$_5$ and should be used as the input nodes in the ANFIS model. A significance level of 0.05 ($p$-value) is assumed to determine which water quality parameters should be included during the model development. In case of Ping-lin station, three explanatory variables, pH, T, and DO, are significant and selected as the input variables. Further examination of the rationality of selection process based on the t-test is conducted. The variation of BOD$_5$ depends on T, pH, microorganisms/organic material, and enzymes activities, and then the organisms and enzyme activities can be further inferred by both T and pH (Karube et al. 1977; Heddam et al. 2016). DO is the maximum oxygen supply in the water and COD represents all the organic substance in the water, so both are related to the changes of BOD (Jouanneau et al. 2014). Accordingly, the BOD$_5$ can be rationally explained by T, pH, and DO/COD, which were all selected as the inputs for the development of ANFIS model. The developed ANFIS model, denoted as ANFIS(I), is given in the form as follows:

$$\text{ANFIS(I): } \text{BOD}_5(t) = f(\text{pH}(t), T(t), \text{DO}(t))$$

After determining the appropriate input variables, ANFIS models with different numbers of MFs are tested to determine the optimal architecture of the system. The RMSE values are used as a criterion to evaluate the performance of the potential models. In order to prevent overfitting, two to eight MFs are evaluated to identify the optimal architecture for each sampling station. The number of MFs increases one by one, and the model with minimum RMSE value for the training data set and significant increasing of RMSE value for the testing data set is accepted as the optimum among potential models. In the case of Ping-lin station, the performance of ANFIS models are sensitive to the number of MFs in their associated architectures (Figure 3(a)). It can be observed that when the number of MFs increases from four to five, the RMSE value for the training data set improves insignificantly, but the increase of RMSE value for the testing set is apparent. This situation implies overfitting is encountered. Overfitting is suggested when the error on the training set is driven to a very small value, while for the testing data presented to the model the error is large. That means the model has memorized the training examples, but fails to generalize to new situations. Hence, the optimal number of MFs is set equal to four for Ping-lin station.

The selection from various MFs, i.e. Gaussian, bell-shaped, trapezoidal-shaped, and triangular-shaped functions, is also evaluated on the basis of RMSE for the training data set. The result of Ping-lin station shows that applying bell-shaped MF could result in better model performance than others (Figure 3(b)). Therefore, in the case of Ping-lin station, an ANFIS(I) model with three input variables, four MFs, and bell-shaped MF could give the optimal result. The R values for the training and testing data sets are 0.91 and 0.86, respectively. The respective values of RMSE, MAPE, and MAE are 0.13, 21.66, and 0.09 for the training data set, and 0.17, 27.17, and 0.09 for the testing data set (Table 2). Figure 4(a) displays the entire time series (including training and testing data sets) of simulated and measured BOD$_5$ at Ping-lin station. A very close pattern of variation among the simulated and measured values is observed. The ANFIS(I) model not only depicts the trend of the data, but also describes the extreme values very well.

The identifications of ANFIS(I) models for other sampling stations were followed the same procedures.
described previously for Ping-lin station. The optimized architecture of ANFIS(I) model at each sampling station and its associated performance are shown in Table 2. According to the results of t-test, the number of input variables varies from two to four. The water quality parameters of pH, T, and DO/COD are included in the identified models for most sampling stations, and the parameter of AN is only included in the model at Jin-gua-liau Creek station. The applied number of MFs varies from three to four, and the usage of bell-shaped function as the MF for all stations resulted in better model performance than other MFs. For the training data sets, the averaged R value for all sampling stations is about 0.9, and for the testing data sets, the R value for each sampling stations is higher than 0.85, except Shui-yuan Bridge station. Regarding the testing data sets, ANFIS(I) model at Da-lin Bridge station has the best performance with highest R, lowest RMSE, and lowest MAE values of 0.91, 0.11, and 0.07, respectively. ANFIS(I) model at Shui-yuan Bridge station has the worst performance with lowest R, highest RMSE,
and highest MAE values of 0.75, 0.25, and 0.16, respectively. Figure 4 displays the entire time series for the simulated and measured BOD$_5$ at all sampling stations. It is clear from these figures that the simulated values of each identified ANFIS(I) model are close to the corresponding measured values. Not only are the trends of the data but also the extreme values well matched. These results provide for the good predictive capability of the identified ANFIS(I) models. The scatter plot of the measured versus simulated values for all sampling stations is shown in Figure 5(a). It can be seen that most points fall along the 1:1 straight line and the fit is quite good. Some scatter points are shown at low BOD$_5$ value and this could be caused by the insensitivity of low BOD$_5$ to the input data. Other factors (or processes) rather than the selected water quality parameters might exist to affect the BOD$_5$ concentration at the low level. The performance of ANFIS(I) model at higher BOD concentration than that at lower BOD concentration indicates its strength of application in a predictive or management framework. Plots of the residuals versus simulated values can be more informative regarding model fitting to a data set. If the residuals appear to behave randomly, it suggests that the model fits the data well. On the other hand, if non-random distribution is evident in the residuals, the model does not fit the data adequately (Singh et al. 2007). The scatter plot of residuals and simulated BOD$_5$ in ANFIS is examined (not shown), and the distributions of residuals at all stations are random and confirm the good fitting of identified models.

![Figure 4](https://iwaponline.com/wst/article-pdf/76/7/1739/450519/wst076071739.pdf)
In order to compare the results obtained from ANFIS(I) models and verify the selection of input variables based on the significance level of the $t$-test, all of the water quality parameters, i.e. five or seven, are used as the input variables to develop the ANFIS models. The newly identified ANFIS model at each sampling station is denoted as ANFIS(II).

Figure 5 | Scatter plot of measured and simulated BOD$_5$ of (a) ANFIS(I), (b) ANFIS(II), (c) MLR(I), (d) ANN(I), (e) MLR(II), and (f) ANN(II) at all sampling stations.
and given in the form as follows:

\[
\text{ANFIS(II)}: \quad \text{BOD}_5(t) = f(\text{pH}(t), T(t), \text{TU}(t), \text{AN}(t), \text{DO}(t)).
\]

or

\[
\text{ANFIS(II)}: \quad \text{BOD}_5(t) = f(\text{pH}(t), T(t), \text{TU}(t), \text{AN}(t), \text{DO}(t), \text{SS}(t), \text{COD}(t))
\]

Again, once the input variables are determined, ANFIS(II) models with a different number of MFs and different MFs are tested to determine the optimal architecture of the system on the basis of RMSE. In case of Ping-lin station, the ANFIS(II) model with five input variables, five MFs, and the bell-shaped MF could give the optimal result. The R values for the training and testing data sets are 0.95 and 0.91, respectively. The respective values of RMSE, MAPE, and MAE are 0.10, 15.26, and 0.06 for the training data set, and 0.15, 25.32, and 0.08 for the testing data set. Compared to the ANFIS(I) model at Ping-lin station, a slight increase of R value and decreases of RMSE, MAPE, and MAE values indicate a trivial improvement of predictive capability of the selected ANFIS model to the BOD\textsubscript{5} data set. The performance of ANFIS(I) model is almost equivalent to the ANFIS(II) model. This comparison demonstrates the selection of input variables based on the t-test of MLR is appropriate, and it is unnecessary to include all of the available parameters during the model development.

The identifications of ANFIS(II) models for other sampling stations followed the same procedures described previously for Ping-lin station. The optimized architecture of ANFIS(II) model at each sampling station and its associated performance are shown in Table 2. Comparing the ANFIS(I) and (II) models at all stations, except Shu-yuan Bridge station, less than 0.05 improvement of R value and 0.1 decrease of RMSE value show that the predictive capability of both models are almost the same. At Shu-yuan Bridge station, only two parameters are used as the input variables in the ANFIS(I) model and the underlying relationship between other parameters and output might not be detected by the t-test. When all of the available parameters are used to develop the model, the unknown relationships are embedded in the model and cause the R value to increase from 0.75 to 0.84. The simulated time series of BOD\textsubscript{5} by ANFIS(II) are almost identical to those by ANFIS(I) (Figure 4) and show the good predictive capability of the identified models. The scatter plot of the measured versus simulated values is shown in Figure 5(b). The points are also concentrated on the 1:1 straight line which implies the good fitness of data. Again, the above results imply that the t-test of MLR can be applied to the selection of input data during the model development and the identified models can produce accurate simulated results.

**COMPARISON TO ANNS AND MLR MODELS AND PREVIOUS LITERATURE**

In this section, the capabilities of proposed ANFIS models are compared to MLR and ANNs models. The entire data set divided into the training and testing sets during the ANFIS identifications is also used for the MLR and ANNs identifications. Two different sets of input variables, one consisting of only significant parameters based on the results of t-test and the other consisting of all available parameters, are used to develop both (I) and (II) models for MLR and ANNs. For ANNs identifications, the number of hidden layers, the number of nodes in the hidden layer, the selection of activation function in each layer, the learning rate, and the momentum are tested to identify the optimal ANNs architecture.

The optimized architectures of MLR and ANNs models at each sampling station and their associated performance of testing data set are shown in Table 3. Comparing the MLR(I) with MLR(II) models and ANNs(I) with ANNs(II) models, a slight improvement of performance in terms of R, RMSE, MAPE, and MAE values indicate that the implementation of the t-test of MLR during the model training is reliable and can be applied in the model development to avoid the tedious work of input selection. Comparing the MLR(I) with ANNs(I) models and MLR(II) with ANNs(II) models, an increase of R value and decreases of RMSE, MAPE, and MAE values show that the ANNs method has the better predictive capability than MLR method. The results of MLR and ANNs are also compared to the identified ANFIS(I) and (II) models at each station (Tables 2 and 3), and a significant improvement of performance in terms of R, RMSE, MAPE, and MAE values confirms that the ANFIS method has the best predictive capability to BOD\textsubscript{5} among the selected methods.

Figure 6 display the entire time series of the simulated and measured BOD\textsubscript{5} by MLR(I) and ANNs(I) models at all sampling stations. From these figures, it is observed that MLR(I) models can only depict the trend of time series data of BOD\textsubscript{5}, but the extreme concentrations, especially high values, are not well described. ANNs(I) models not only depicts the trend of the data, but also
roughly describes the extreme values which cannot be done by the MLR(I) models. The simulated values obtained by ANNs(I) have a closer pattern than those obtained by MLR(I) and are suggested for the better predictive capability of the ANNs (grey line vs. blue line in Figure 6; please refer to the online version of this paper to see this figure in color: http://dx.doi.org/10.2166/wst.2017.359). Compared to the ANFIS models (Figures 4 and 6), either ANFIS(I) or (II) models not only depicts the trend of the data, but also describes the extreme values better than both MLR and ANNs models. The scatter plots of measured and simulated BOD$_5$ at all sampling stations for the MLR and ANNs models are shown in Figure 5(c)–5(f).

From Figure 5(c) and 5(e), it can be seen that the MLR estimates are far away from the corresponding measured values, and many scatter points fall below the 1:1 straight line, especially with the higher measured values. This phenomenon indicates that the MLR models underestimate the BOD$_5$ values, and this situation is more severe when measured concentrations are higher. Compared to the MLR, the ANNs estimates are more concentrated on the 1:1 straight line, and fewer points are underestimated, particularly the higher values (Figure 5(d) and 5(f)). Comparing Figure 5(a)–5(f), the estimates of ANFIS models are closer to the corresponding measured values than those of MLR and ANNs models and fewest points are underestimated at the higher values. These results suggest that the ANFIS model with a more complex architecture is more suitable than ANNs or MLR models to describe the nonlinear relationship within the variables caused by the complexity of aquatic systems and to produce the best fit of the measured BOD$_5$ concentrations. The better performance of ANFIS model at higher BOD concentration further reveals its capability in the development of management strategies for the protection of water resources.

Although the comparison of computational time (or efficiency) during the training process between ANFIS and ANNs is not the scope of this study, the examination of computational time is conducted and the time difference for model training between two models is similar. Accordingly, the increase of accuracy in ANFIS model might be the consequence of including fuzzy logic theory during the model development. The tradeoff in terms of computational time and accuracy is obvious between ANFIS (ANNs) and MLR. The development of MLR model is very fast, but the predictive accuracy is suspicious.

In the last decade, some studies have demonstrated the implementation of ANNs or MLR modeling to predict DO/BOD concentrations in watershed or river systems (Soyupak et al. 2003; Oliveira-Esquerrre et al. 2004; Singh et al. 2009; Rankovic et al. 2010; Ay & Kisi 2012). The performance of their models showed good predictive abilities in term of R value higher than 0.8. In this study, the R values obtained by ANFIS(I) for the training and testing sets at eight sampling stations ranged from 0.87–0.94 and 0.75–0.91, respectively, and the results are comparative to those reported in the previous literature.

| Table 3 | The basic structure and performance of identified MLR(I), MLR(II), ANN(I), and ANN(II) models for the testing data set at each sampling station |
| Station | Input (I)/(II) | Method | Number of nodes (I)/(II) | R (I)/(II) | RMSE (I)/(II) | MAPE (I)/(II) | MAE (I)/(II) |
| Ping-lin | pH, T, DO/ALL | MLR – | | 0.52/0.53 | 0.30/0.30 | 117.88/114.87 | 0.26/0.25 |
| | | ANN 9/12 | | 0.68/0.72 | 0.27/0.27 | 78.27/61.30 | 0.25/0.23 |
| Dai-yu-ku Creek | pH, T, DO/ALL | MLR – | | 0.45/0.49 | 0.31/0.30 | 95.33/92.49 | 0.25/0.24 |
| | | ANN 13/18 | | 0.53/0.58 | 0.27/0.24 | 90.38/84.27 | 0.21/0.18 |
| Jin-gua-liau Creek | pH, T, AN, DO/ALL | MLR – | | 0.16/0.17 | 0.53/0.53 | 112.04/110.17 | 0.52/0.50 |
| | | ANN 15/25 | | 0.25/0.36 | 0.48/0.46 | 107.11/99.39 | 0.43/0.40 |
| Kuo-lai | pH, T, COD/ALL | MLR – | | 0.11/0.13 | 0.35/0.34 | 91.64/82.14 | 0.32/0.32 |
| | | ANN 10/13 | | 0.51/0.53 | 0.22/0.18 | 43.28/42.81 | 0.27/0.21 |
| Bi-hu | DO, COD/ALL | MLR – | | 0.15/0.20 | 0.46/0.46 | 93.45/92.41 | 0.28/0.25 |
| | | ANN 15/20 | | 0.44/0.51 | 0.33/0.31 | 72.43/71.58 | 0.22/0.21 |
| Da-lin Bridge | pH, T, DO, COD/ALL | MLR – | | 0.21/0.22 | 0.34/0.34 | 81.69/81.28 | 0.19/0.18 |
| | | ANN 10/16 | | 0.52/0.54 | 0.26/0.25 | 78.16/68.32 | 0.18/0.17 |
| Shui-yuan Bridge | T, COD/ALL | MLR – | | 0.23/0.26 | 0.40/0.39 | 109.65/105.96 | 0.50/0.40 |
| | | ANN 15/23 | | 0.48/0.56 | 0.31/0.26 | 66.37/55.31 | 0.24/0.24 |
| Huang-ju-pi-liao | pH, T, COD/ALL | MLR – | | 0.37/0.39 | 0.41/0.41 | 65.01/62.84 | 0.35/0.33 |
| | | ANN 8/10 | | 0.48/0.50 | 0.31/0.22 | 53.46/48.27 | 0.32/0.31 |
CONCLUSION

Modeling water quality parameters is a very important aspect in the analysis of any aquatic system. Despite the numerous numerical or surrogate models available, the research for improving the effectiveness and accuracy of forecasting models has never stopped. In this study, ANFIS is selected and applied to model BOD\textsubscript{5} concentrations at the upstream catchment of Feitsui Reservoir, Taiwan. The water quality data are collected at eight sampling stations from 1987–2012 and eight different water quality parameters, pH, T, TU, AN, DO, SS, COD, and BOD\textsubscript{5}, are used to develop the ANFIS models. The significant level of t-test obtained from MLR is applied to discover the appropriate input variables in the ANFIS model. The potential architectures of ANFIS models with different number of MFs and the MFs are tested, and the performance of testing data set is carefully examined to avoid overfitting.

The result of t-test shows that the water quality parameters of pH, T, DO/COD are the most significant parameters to the BOD\textsubscript{5} for most of sampling stations and are used as the input variables for the model training processes. The examination of underlying processes of BOD\textsubscript{5} in an aquatic system reported in previous literature (Karube et al. 1977; Jouanneau et al. 2014; Heddam et al. 2016) confirms that the selection process of water quality parameters based on the t-test is mathematical sensible, not a coincidence, and can be extended to other kinds of neural-based models to avoid the tedious work of parameter

Figure 6 | Simulated and measured BOD\textsubscript{5} of MLR(I) and ANN(I) models at (a) Ping-lin, (b) Dai-ya-ku Creek, (c) Jin-gua-liau Creek, (d) Kuo-lai, (e) Bi-hu, (f) Da-lin Bridge, (g) Shui-yuan Bridge, (h) Huang-ju-pl-iaio.
selection. The simulated BOD$_5$ concentrations at each sampling station are closer to the corresponding measured values, and both trends and extreme values of the data are well matched. The capability of ANFIS method along with the significant test to the selection of water quality parameter of BOD$_5$ is demonstrated. Compared to the MLR and ANNs models, the increase of accuracy in ANFIS model might be the consequence of including fuzzy logic theory in the model development. The predictive ability of the ANFIS model, especially at higher BOD concentration, indicates its strength of application in the framework of water resources management.

Although the ANFIS model, similar to the ANNs model, can be seen as a black box and the physical mechanisms of the natural system are not explicitly represented, its powerful capability of pattern recognition could be applied as an effective tool for the computation of river water quality. ANFIS models can efficiently deal with numerous and complex input-output patterns because of their nonlinearity and linguistic uncertainty. They have great ability to provide effective and accurate predictions for complicated physical, chemical, and biological processes of BOD$_5$ in aquatic systems. ANFIS can be seen as a predictive alternative to traditional modeling techniques and extended to other parameters of water quality. Future research on the issues of climate change and applications to other watersheds may be required to strengthen these conclusions. In this study, the pollution sources of BOD$_5$ and wastewater discharges are not considered in the model development. The establishment of relations between riverine water qualities and amount of fertilizer by farmers) could also be the wastewater discharge/pollution sources (e.g. the timing and amount of fertilizer by farmers) could also be the wastewater discharge/pollution sources (e.g. the timing and amount of fertilizer by farmers) could also be the wastewater discharge/pollution sources (e.g. the timing and amount of fertilizer by farmers). ACKNOWLEDGEMENTS

This material is based on the work supported by National Science Council (NSC) under award NSC 102–2116–M–019–001. The data supported by Taipei Feitsui Reservoir Administration and Taipei Water Management Committee are gratefully acknowledged.

REFERENCES


First received 6 September 2016; accepted in revised form 2 June 2017. Available online 16 June 2017