WCIT-2010

Prediction of primary treatment effluent parameters by Fuzzy Inference System (FIS) approach

Esra Yel, Sukran Yalpir

* Selcuk Univ., Environmental Eng. Department, 42075 Campus KONYA-TURKEY
b Selcuk Univ., Geomatics Eng. Department, 42075 Campus KONYA-TURKEY

Abstract

A fuzzy-logic-based diagnosis system was developed to determine the primary treatment effluent quality in a municipal wastewater treatment plant (MWTP). The measured data of variables were implemented into the Fuzzy Inference System (FIS) with Mamdani’s method. The fuzzy control rule base was shaped to define essential quality parameters monitored as pH, COD, BOD and SS outputs.

The output approximations to real data remained in an acceptable range for a MWTP performance (89-96%). The averages and standard deviations of the model were also approximated closely as 93-98% and 89-97%, respectively. The resulting configuration proved a good modeling approach for MWTP effluent quality prediction.

Keywords: Fuzzy Inference System, Wastewater, Expert Knowledge, Mamdani.

1. Introduction

Recent rapid urban development in and around residential areas in the metropolitan regions puts a heavy burden on the environment which is often undervalued against economic development and industrialization, especially in developing countries. The intensity of urbanization can be measured by different parameters where establishing and understanding the relationship between waste management and the urbanization may help future planning and efforts to alleviate wastewater problems in the region. The effect of the developmental density restrictions associated with water supply protection on land prices was studied. The same approach can be integrated to waste management which may have potential to affect urbanization and land prices. Population growth and land use changes will likely increase the pressure on water supply and water quality. This, in turn, requires a continued control of watershed management and protection practices to support a sustainable wastewater management, such as protecting the watersheds, reducing risks to humans and the environment and increasing reuse opportunities. At this point, the performance assessment in municipal wastewater treatment plants (MWTP) can be used to provide baseline management status for discharge to receiving media which may be located in a watershed.

Variations in raw wastewater (WW) composition, strength and flow rate make MWTP operational control difficult. When this variability is combined with the complex nature of the treatment processes, the system becomes more complicated and modeling becomes a need for proper operational control. In order to follow the MWTP performance during the operation, effluent monitoring is not sufficient, where the operator has to control the system...
and to take necessary precautions upon the influent WW characteristics before any problem in treatment performance has arisen. Prediction of the effluent quality can only be achieved by modeling via mathematical modeling methods. Recent studies based on Artificial Intelligence (AI) have been applied on modeling of MWTP output parameters [1].

It is clearly mentioned by Mingzhi et al [2] that in spite of some successful practical applications, there is still no all-inclusive procedure or method to design such intelligent controllers by far because of its semi-empirical nature. It is also reported in the literature that, consideration of statistical principles in AI model building process may improve modeling performance [3, 4].

Adaptive Neural Fuzzy Inference Systems (ANFIS) and Artificial Neural Network (ANN) are the artificial intelligence techniques widely preferred for modeling MWTP parameters [5-15], however preference of Fuzzy Inference System (FIS) for the same purpose has not been observed in the literature.

Fuzzy logic is a way to implement expert knowledge (EK) to establish an advanced control on various treatment processes which lacks in application to primary treatment units in a wastewater treatment plant as a whole.

Fuzzy inference systems have found a wide range of industrial and commercial control applications that require analysis of uncertain and imprecise information [16]. Recently, prevailing research efforts on fuzzy logic control have been devoted to model-based fuzzy control systems that guarantee not only stability but also performance of closed-loop fuzzy control systems [17]. The fuzzy inference system consists of membership functions for state and control variables, production rules prepared from information provided by experienced operators and a fuzzy inference engine [18]. Professional experience is a critical contribution in the evaluation of MWTP systems and in the solution of operational problems during the plant operation.

In this study a fuzzy-logic-based diagnosis system was developed for the determination of the primary treatment effluent quality. The information of measured variables and the EK were implemented into the fuzzy inference system (FIS) with Mamdani’s method by means of a fuzzy-based rule structure.

2. Methodology

2.1. Municipal Wastewater Treatment Plant

Data was taken from the database of a MWTP designed for 1.4 million population equivalent in Turkey. In the model structure the primary treatment units are taken into consideration (Figure 1). To construct the model structure, totally 5 critical wastewater quality parameters (indicated in Table 1) were selected as input variables. The output of the FIS model includes pH, chemical oxygen demand (COD), biochemical oxygen demand (BOD) and suspended solids (SS). 52 non-consecutive test data were selected from the yearly measurements of the MWTP. Selection was based on the seasonal variations which have influence on effluent parameters.

![Figure 1. Primary Treatment System in MWTP](image)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Maximum</th>
<th>Average</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>9.20</td>
<td>8.10</td>
<td>7.20</td>
</tr>
<tr>
<td>Temperature °C</td>
<td>24.46</td>
<td>20.27</td>
<td>15.60</td>
</tr>
<tr>
<td>COD (mg/L)</td>
<td>785</td>
<td>662</td>
<td>424</td>
</tr>
<tr>
<td>BOD (mg/L)</td>
<td>430</td>
<td>334</td>
<td>200</td>
</tr>
<tr>
<td>SS (mg/L)</td>
<td>488</td>
<td>333</td>
<td>171</td>
</tr>
<tr>
<td>pH</td>
<td>8.80</td>
<td>7.98</td>
<td>7.10</td>
</tr>
<tr>
<td>COD (mg/L)</td>
<td>652</td>
<td>512</td>
<td>386</td>
</tr>
<tr>
<td>BOD (mg/L)</td>
<td>340</td>
<td>265</td>
<td>190</td>
</tr>
</tbody>
</table>

Table 1. Maximum, Average and Minimum values of the MWTP Parameters
2.2 Model Structure and Configuration

The topology of the network composed of 4 layers in the development of a fuzzy system (Figure 2).

**Input, General Information Base Unit:** involves the input variables affecting the considered event and all the information related to these variables. The “general database” term is used due to the possibility of having information in numerical and/or text formats (pH, Temperature, COD, BOD, SS in this study). The model configuration given in Table 2 was structured as a result of EK implementation. The selections given in Table 2 provide a basis for the development of rule base.

**The Fuzzy Maker:** is a processor assigning numerical input values to membership grades in fuzzy sets characterized with text (Common membership functions are triangular, bell curved and trapezoidal functions. The triangular function was selected as the main membership function of this study. The last memberships were transformed into trapezoidal function for some variables; if their ends were open i.e. if they were greater than the value of the last interval).

**Table 2.** The Model Configuration

<table>
<thead>
<tr>
<th>Input</th>
<th>pH</th>
<th>COD</th>
<th>BOD</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Temperature</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>COD</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BOD</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fuzzy Rule Base Unit:** contains all of the rules writeable in logical IF – THEN expression connecting input variables to output variables in the database. In writing these rules, all possible intermediate (fuzzy set) connections between inputs and outputs are taken into consideration. The fuzzy system can be applied in two ways each having different rules. The rule base was formed after assignment of the memberships. In the rule base, totally 80 rules (19, 40 and 21 different rules defined in accordance with EK for output pH, COD and SS, respectively; the same rules were used for COD and BOD) were written in “If x is A and y is B, then z is C” format, e.g., “IF pH is Normal AND Temperature is low AND …. THEN pH is high”.

**Fuzzy Inference Motor Unit:** is a mechanism covering the group of processes providing the single output behavior of the system by gathering the separate relations built between the input and output fuzzy sets in the fuzzy rule base. This motor is used to determine what kind of an output will be obtained as a result of the input of the whole system by collecting all the rule inferences together.

**Defuzzifier:** transforms the fuzzy inference solutions obtained as a result of fuzzy processes into definite numerical output values. The results of the rules were combined and defuzzified via centroid method.

**The Output Unit:** expresses the group of the output values obtained at the end of the interaction performed between information and fuzzy rule bases by the help of the fuzzy inference motor (output pH, COD, BOD, SS)
2.3. Statistical evaluation

The model results were statistically analyzed using Mean Absolute Percentage Error (MAPE) (1).

\[
MAPE\% = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{x_p - x_i}{x_p} \right| \times 100\%
\]

(1)

where, \(x_p\) is measured value, \(x_i\) is value of the FIS structure, \(i\): \{1,2,3...n\}, \(n\) is the total number of data in the dataset. In evaluating the model performance, standard deviations (SD) and averages of both the real measurements and the model results were calculated. Model statistics were expected to close to real data SD and averages. This was calculated with the approximation percentages of the average (AA\%) and SD (SDA\%).

3. Results

To develop estimation model that can provide accurate predictions of the effluent parameters, a preprocessing may be helpful in input data selection as well as engineering judgment. This procedure may include selection of wide range of parameter values, involvement of all seasons’ data etc.

Application results for FIS are presented in Figure 3. As the model values and measured values were compared for all the parameters, the model performed very close to measured values in predicting the extreme points dropping far from the average, and performed well in the prediction of the points in the general distribution. Highly fitting predictions were obtained in many data points for all parameters. Despite some inconsistencies, the patterns of predicted and measured values were parallel. Considering the fluctuating characteristics of the influent wastewater and primary treatment units, the prediction performance for each parameter (Figure 3) were evaluated separately. pH range in wastewater treatment plants is narrow in nature (generally 6 to 9), therefore, fluctuations occur in the range of very low scales. The absence or low occurrence of extremely high/low pH values limits the model’s predicting ability in case of extreme values for the models applying training and testing algorithms without EK, whereas, FIS includes pH values lower than 6 and higher than 9 depending upon EK.

Primary treatment is generally applied for solids removal in screening, grid removal and primary sedimentation units, organic matter (COD and BOD) removal is not the main target. A 30-60 % removal in SS is usually obtained.
in the primary treatment and the rest is removed in the secondary sedimentation tank, therefore, the values can fluctuate and the system can tolerate these variations. In case of temporary shock loads and dilutions in the influent values, it is inevitable to experience fluctuations through the whole plant units. For COD removal is not a priority target in the primary treatment, it is not appropriate to expect a close relation in the effluent. About 25 to 40% COD-BOD removal is achieved. The fluctuations in these parameters remained in the acceptable range for both the model and the real data.

The average and SD values achieved in the outputs were close to those of measured data (Table 3). As MWTP produce a highly fluctuating data due to varying input wastewater characteristics and process efficiency, monitored parameters in the outlet flow follow a pattern with a high SD, therefore, the model performance has been evaluated with approximation approach in the average and SD between the measured and calculated data. The model produced a SD with an approximation at 89-97% to the measured data for all output parameters. The prediction level was even better with the average value with the average value approximation at 93-98%, moreover, the wellness of the approach was substantiated with a 4-11% range of MAPE which assured the high performance of the model indicating that deviation produced by the model coincided with real data at most of the data points.

As the same data was modelled with ANFIS approach in the authors’ previous work [10], the model tended to linearize its output in all the parameters where the SD values of pH, COD, BOD and SS were obtained at lower level as 0.15, 28, 22 and 16, respectively. The approximation stayed at a considerably low level as 46, 52, 67 and 49% compared to FIS approach which proved its high performance as a control system feasible in MWTP indicating that
an EK based modelling approach can approximate at a much efficient level than approaches with learning process such as ANFIS, possessing limitations such as using a part of the available data for learning and testing purposes.

4. Conclusions

The resulting configuration is a good modeling approach for wastewater treatment plant effluent quality prediction. The proposed FIS approach is effective and reliable for extracting features from input data in combination with EK. The high SD’s originated from varying influent wastewater characteristics and process efficiencies were approximated at more than 89% level in the proposed model. The calculated errors remained at 4-11% level, which falls within the acceptable limits as the analytical error intervals are much higher in the measurements of these parameters (generally 10-20%). As the model configured in FIS structure involves expertise more as compared to other AI techniques, the complex nature of the parameters and the treatment units fit this structure better. Moreover, the model can respond input values that have not been involved in the initial data. Therefore, this structure does not require the separation of training and test data set.

References