Evaluation of Biological Oxygen Demand from Sugar Industry Effluents - An Neuro Fuzzy Approach

S. Akilandeswari*, A. Julie.
Department of Physics, Annamalai University, Annamalainagar 608 002, Tamilnadu, India
E-mail: akilaphy2010@gmail.com

Received 26 November 2012; accepted 19 December 2012

Abstract

Several intelligence models are being increasingly employed for the prediction of BOD (Biological oxygen demand) from the industrial effluents. These models can be described as mathematical methodologies which explain input and output relations irrespective of the process and without the need for making assumptions considering the nature of the relations. The New adaptive neural fuzzy inference system (ANFIS) is an intelligent technique, which can be successfully used to predict the values of BOD from the known characteristic parameters of the effluent. The BOD values of effluents from a sugar industry are estimated by both ANFIS and statistical modeling and the results are compared with observed values. The main purpose of this study is to investigate the abilities of an ANFIS model to improve the accuracy of the biological oxygen demand (BOD) estimation.

Keywords: Industrial effluent, BOD, ANFIS, SPSS.

1. Introduction:
Sugar industry plays an important role in the economy of India by way of farming and creation of employment. Effluent from sugar industry with its high biological oxygen demand (BOD) rapidly deplete available oxygen supply when discharged into water bodies endangering aquatic life and also creates foul smelling hydrogen sulfide, which in turn can precipitate iron and any dissolved salts, turning the water black and highly toxic for aquatic life. Biological oxygen demand is one of the major pollution parameter of effluents. It is an approximate measure of the amount of biochemical degradable organic matter present in a water sample. It is defined by the amount of oxygen required for the aerobic micro organism present in the sample to oxidize the organic matter to a stable organic form [1]. The BOD test requires significant time and commitment for preparation and analysis. The entire process requires five days, with data collection and evaluation occurring on the last day. A test is used to measure the amount of oxygen consumed by these organisms during a specified period of time (usually 5 days at 20°C). The difference in initial DO readings (prior to incubation) and final DO readings (after 5 days of incubation) is used to determine the initial BOD concentration of the sample. This is referred to as a BOD₅ measurement [2]. Many BOD investigations have been made to decrease the time required for determining this parameter [3, 4, 5, 6]. In this study, Adaptive Neuro Fuzzy Inference System (ANFIS) is a potential approach, for water quality modeling introduced by Jang [7] which is based on the first-order Sugeno fuzzy type [8]. In a very simplistic analysis, it can be said that the ANFIS is a first-order Sugeno fuzzy system, characterized by a set of structured fuzzy rules of the type IF (observation) THEN (actuation) that is ‘adapted’ in order to be able to incorporate the learning capacities of the neural network. ANFIS is so designed that it can be used in modeling, decision making and process control applications. The main purpose of this study is to analyze and compare the performance of ANFIS and Statistical models in predicting biochemical oxygen demand from sugar industry effluent.

2. Materials and methods:
For the present study the effluent samples were collected from sugar industry at the sources and were analyze the parameter in the laboratory. Physico-chemical parameters were analyzed according to APHA [9]. In the present study both the statistical multiple regression analysis and ANFIS modeling are employed to evaluate the BOD.

2.1 ANFIS Modeling:
Adaptive Neuro Fuzzy Inferences System (ANFIS) approach was employed in this study. The ANFIS architecture consists of fuzzification layer, inferences process, defuzzification layer, and summation as final
Fig 1. Anfis architecture with two inputs and an output output layer. Typical architecture of ANFIS is shown by Figure 1.

The process flows from layer 1 to layer 5. It is started by giving a number of sets of crisp values as input to be fuzzyfied in layer 1, passing through inference process in layer 2 and 3 where rules applied, calculating output for each corresponding rules in layer 4 and then in layer 5 all outputs from layer 4 are summed up to get one final output. The main objective of the ANFIS is to determine the optimum values of the equivalent fuzzy inference system parameters by applying a learning algorithm using input-output data sets. ANFIS generates automatically the fuzzy rules and selects the rules with maximum firing strength. The parameter optimization is done in such a way during training session that the error between the target and the actual output is minimized. Parameters are optimized by hybrid algorithm which combination of least square estimate and gradient descent method. The parameters to be optimized in ANFIS are the premise parameters which describe the shape of the membership functions, and the consequent parameters which describe the overall output of the system. The optimum parameters obtained are then used in testing session to calculate the prediction. In the present work, programme is written to work from the command line, using the Fuzzy logic toolbox supported in MATLAB version 5.3. The radius and number of epochs are so selected that the training root mean square error (Trn RMSE) and the check root mean square error (Chk RMSE) minimum. The accuracy of the predicted values of BOD can be checked by calculating the Average percentage error (APE) using the relation [7].

\[
APE = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{|BOD_{(obs)} - BOD_{(pred)}|}{BOD_{(obs)}} \right) \times 100\%
\]

Where, n is number of data pairs

2.2 Statistical analysis:

When many independent parameters have influence on one dependent parameter, the statistical multiple regressions analysis can be effectively utilized in the prediction of the dependent parameter [10]. Multiple linear regression (MLR) was applied to predict the dependent variable. A MLR model takes the form \( Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \ldots \)

Where \( Y \) is the predicting variable, \( x_1, x_2 \ldots \) are independent variables, with parameters \( \beta_0, \beta_1, \beta_2 \ldots \ldots \) are
regression coefficients. Statistical analysis was carried out using SPSS version 14.0. Data obtained from statistical analysis are compared with experimental values. Chi-squared test is a statistical tool commonly used to compare observed data with data we would expect to obtain according to a specific hypothesis. Chi-square is the sum of the squared difference between observed \( (O) \) and the expected \( (E) \) data divided by the expected data in all possible categories. The Chi-square test was calculated in both cases by using the following relation.

\[
\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}
\]

### 3. Results and discussion:

The training data of physico-chemical characteristics of sugar industry effluents are used in the present work is shown in Table 1, and the predicted values of BOD using statistical and ANFIS modeling and the observed characteristic parameters are shown in Table 2. Average percentage error (APE), chi-square \( (\chi^2) \), and Worst case error (WE) were calculated. By comparing the observed values of BOD with the predicted BOD from statistical model and ANFIS model, the experimental result shows that the predicted BOD from statistical model shows more deviation than observed value.

From the tabel,2, the predicted ANFIS model shows low average percentage error (APE), low Chi-Sqr value, and low worst case error than the statistical model. It can be observed that the results given by the ANFIS model are superior to the statistical model.

The data furnished in table 1, graphically presented in fig 2. From the graph it is clear that ANFIS model gives higher accuracy than the statistical model, and the ANFIS model is very good agreement with the observed value.

### 4. Conclusion:

In this paper, we compared two models, statistical and adaptive neuro-fuzzy inference system, for the estimation of biological oxygen demand of sugar industry effluent. The experimental results show that the adaptive neuro-fuzzy inference system model provided lower APE, lower chi-sqr value, and lower worst case error than the statistical model. On the basis of the results, the ANFIS technique was found to be superior to the statistical technique. This result may be applied to automate BOD estimations.

### References

7. Roger Jang, J. (May 1993). *ANFIS : Adaptive-

Source of support: Nil; Conflict of interest: None declared