Volume : 2 | Issue : 10 | Oct 2013 • ISSN No 2277 - 8160

Research Paper | Physics

Estimation of Biological Oxygen Demand Through Anfis Modeling

Dr.S.Akilandeswari — Assistant professor, Department of physics, Annamalai university, Annamalai nagar,608 002 Tamil nadu. India.

A.Julie — Research scholar, Department of physics, Annamalai university, Annamalai nagar,608 002, Tamil nadu. India.

ABSTRACT

Pollution due to discharge of industrial effluents has become a serious problem in most of the areas of our country. Among the various parameters of the effluents, Biological oxygen demand is one of the important parameter. Typically, finding the BOD requires 5 days, with data collection and evaluation occurring on the last day. Nowadays, several techniques such as, Adaptive Neuro Fuzzy Inference System (ANFIS) and statistical models are employed for developing the predictive models to estimate the effluent parameters of biological oxygen demand from industrial effluents. The main objective of this study is to compare between the predictive ability of the Adaptive Neuro-Fuzzy Inference System (ANFIS) model and statistical model with the observed BOD value. The experimental result shows that the ANFIS model provides a higher accuracy than the statistical model.

KEYWORDS : ANFIS, Industrial effluent, BOD, Statistical

1. Introduction

Industrial activities consume a huge amount of natural water, utilizable resources and energy thereby discharging enormous wastewater to the natural environment. It is of great importance in water quality control that the amount of organic matter present in the system be known and that the quantity of oxygen required for its stabilization be determined. Performing the test for BOD requires significant time and commitment for preparation and analysis. This process requires 5 days, with data collection and evaluation occurring on the last day. A test is used to measure the amount of oxygen consumed by aerobic micro 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 BOD_5 measurement [1].Many investigations have been made to decrease the time required for determining this parameter[2,3,4] . Inspired by the capacities of the human brain, artifical intelligence (AI)-based models integrate the specific attributes of various disciplines, such as mathematics, statistics, physics, computer science, and just recently, environmental engineering applications[5]. The new Adaptive Neuro Fuzzy Inference System (ANFIS) model, approach, for predicting effluent quality modeling introduced by Jang [6] which is based on the fuzzy modeling or fuzzy identification, first explored by Takagi and Sugeno[7] has numerous practical applications like control[8,9,10,11] , prediction and inference [12,13].

2. Materials and methods:

2.1. Sample collection:

For the present study the effluent samples were collected from chemical industry at the sources over a period of one year around cuddalore district and were analyze the parameter in the laboratory. Physico-chemical parameters were analyzed according to APHA [14]. In the present study both the statistical multiple regression analysis and ANFIS modeling are employed to evaluate the BOD.

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 [15]. Multiple linear regression (MLR) was applied to predict the dependent variable. A MLR model takes the form $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \ldots$ Where Y is the predicting variable, are independent variables, x_1, x_2, x_3....... with parameters, β_0 β_1, β_2......are regression coefficients. Statistical analysis was carried out using SPSS version 14.0. In the statistical multiple regression analysis, the physico-chemical parameters of the effluent namely $pH, TDS,, Cl^-, So_4^{2-}$ are taken as independent parameters (x_1, x_2, x_3... ...) and the BOD as dependent parameter.

2.3. Adaptive neuro-fuzzy inference system (ANFIS) modeling:

ANFIS is a method based on the input–output data of the system under consideration. Success in obtaining a reliable and robust ANFIS network depends heavily on the choice of process variables involved as well as the available data set and the domain used for training purposes [16]. Basically, a fuzzy inference system is composed of five function blocks [6]:

(i) a rule base containing a number of fuzzy if-then rules.
(ii) a database which defines the membership function of the fuzzy sets used in the fuzzy rules
(iii) a decision-making unit which perform the inference operation on the rules.
(iv) a fuzzification inference which transforms the crisp inputs into degrees of match with linguistic values.
(v) a defuzzification inference which transforms the fuzzy results of the inference into a crisp output.

For simplicity, a fuzzy inference system with two inputs x and y, and one output is assumed. For a first-order Sugeno fuzzy model, a common rule set with two fuzzy if –then rules is defines as follows:

Rules 1: If x is A_1 and y is B_1, then $f_1 = p_1x + q_1y + r_1$............ (1)

Rules 2: If x is A_2 and y is B_2, then $f_2 = p_2x + q_2y + r_2$.......... (2)

Here type-3 fuzzy inference system proposed by Takagi and Sugeno [7] is used. In this inference system the output of each rule is a linear combination of input variables added by a constant term. The final output is the weighted average of each rule's output. The corresponding equivalent ANFIS structure is shown in Fig. 1.

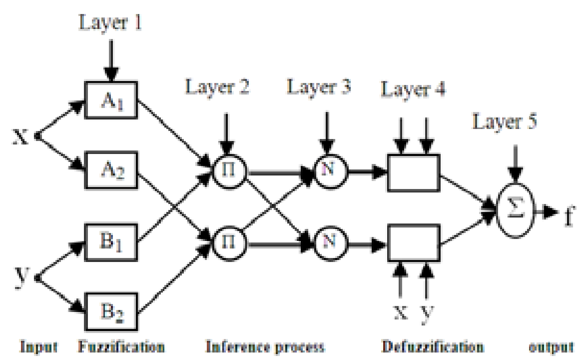


Fig 1. Anfis architecture with two inputs and an output output layer.

Every node in layer 1 is an adaptive node with a node function that may be a Gaussian membership function or any membership functions. Every node in layer 2 is a fixed node labeled Π, representing the firing strength of each rule. Every node in layer 3 is a fixed node labeled N, representing the normalized firing strength of each rule. Every node in Layer 4 is an adaptive node with a node function. The single node in layer 5 is a fixed node labeled Σ, indicating the overall output (Z) as the summation of all incoming signals [18]. In this study, Gaussian membership function is used for the input variable. The hybrid learning algorithm is employed to determine the parameters of Sugeno-type fuzzy inference systems. For a given training dataset, the combination of the least-squares method and the back-propagation gradient descent method is utilized to update FIS membership function parameters.

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.

3.Results and discussion

The Training data of physico-chemical characteristics of chemical 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.

Table.1 Characteristic parameters used as 'Training Data Set' of chemical industry effluent

S.NO	Inputs (selectively used to train ANFIS)				
	pH	TDS (mg /l)	cl^- (mg/l)	So_4^{2-} (mg/l)	BOD (mg /l)
1	7.6	452	88	203	3
2	8.76	408	56	147	2
3	8.2	420	60	23	3
4	7.6	1044	78	29	7
5	7.2	648	363	160	2
6	6.43	484	179	205	15
7	7.59	1860	245	332	2
8	7.32	1072	130	148	3
9	8.14	552	154	126	13
10	7.47	568	142	226	3
11	8.21	528	232	119	4
12	8.4	560	45	113	8
13	8.1	944	591	217	12
14	6.7	808	177	212	7
15	7.42	480	542	186	9
16	7.0	956	227	43	2
17	6.12	689	98	185	5
18	6.08	642	123	168	7
19	7.16	536	326	132	9
20	8.01	936	69	89	3
21	6.06	844	453	186	1
22	6.36	728	331	20	5
23	7.1	540	472	320	9
24	7.23	592	165	26	3
25	7.04	628	725	53	8
26	6.97	2024	86	46	6
27	6.83	1536	130	138	2
28	6.25	1520	304	140	3
29	7.17	2080	162	112	4
30	6.63	832	126	229	2
31	6.71	1476	485	247	4
32	6.76	952	76	46	3
33	7.13	2016	174	29	7
34	6.55	444	183	28	1
35	6.3	892	106	244	3
36	6.53	592	54	46	2
37	7.69	862	98	89	6
38	7.52	712	168	142	4
39	8.02	653	173	96	1
40	6.89	1032	146	214	9

The performance of ANFIS and statistical outputs was evaluated by estimating the Average percentage error, chi-square test, and worst case error value which is defined as, [18]

$$APE = \frac{1}{n}\sum_{i=1}^{n}\frac{|BOD_{(obs)} - BOD_{(Pred)}|}{BOD_{(obs)}} \times 100\%$$

Where, n represents number of data pairs, $BOD_{(obs)}$ represents observed values of BOD, $BOD_{(pred)}$ represents predicted values of BOD.

A very powerful test for testing the significance of the discrepancy between observed and predicted values is the chi-square test of goodness of fit of the model.

$$\chi^2 = \sum_{i=1}^{n}\frac{(O_i - E_i)^2}{E_i}$$

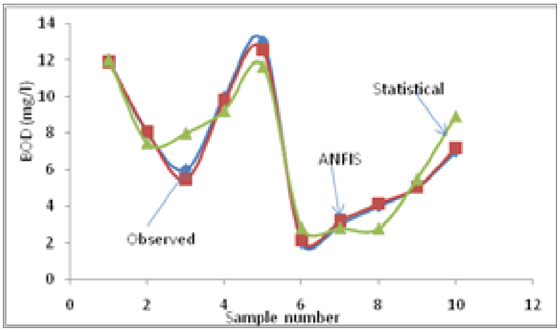
Where ,O_i(i=1,2,3,…….n) is set of observed values and E_i(i=1,2,3,……...n) is the corresponding set of predicted data.

The accuracy in the predicted values of BOD can be also estimated by finding the Worst-case error (WE) using the relation

Table. 2 Characteristic parameters, and) of chemical industry effluents used as' Check Data' the observed and predicted values of BOD.

Sample No	pH	TDS (mg/l)	cl^- (mg/l)	So_4^{2-} (mg/l)	Observed BOD (mg/l)	Predicted BOD		Multiple regression parameters and coefficients
						ANFIS model	Statistical model	
1	7.47	726	654	192	12	11.8982	12.05	
2	6.92	315	532	413	8	8.0747	7.46	
3	6.56	433	542	180	6	5.4598	8.00	
4	7.68	568	426	29	10	9.8078	9.25	
5	8.29	1092	451	227	13	12.5520	11.67	R^2 = 0.901 P = 0.01 F = 11.366
6	7.19	686	89	424	2	2.1266	2.82	β_0 = -12.112 = 1.794 = 0.003 = 0.014 β_4 = -0.003
7	6.08	681	163	102	3	3.2111	2.81	
8	6.49	851	73	108	4	4.1417	2.78	
9	8.54	346	120	143	5	5.0293	5.50	
10	8.03	1536	152	28	7	7.1805	8.95	
					APE =	3.62	17.37	
					WE =	0.5	2.0	
					Chi-sqr=	0.105737	2.008681	

Fig .2 Plot of BOD verses Sample number (with, as inputs)



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 the present study ANFIS and statistical methodology was adopted to model prediction of biological oxygen demand from chemical industry effluents. According to the results, ANFIS and Statistical model could predict the BOD value of industrial effluent. The minimum average percentage error of 3.62 %, worst case error value of 0.5 chi-sqr value of 0.105737 could be achieved using ANFIS. The maximum average percentage error of 17.37 %, worst case error value of 2.0 chi-sqr values of 2.008681 achieved using Statistical model. On the basis of the comparison results, the ANFIS technique was found to be superior to the Statistical technique. It has been shown that application of the proposed knowledge-based fuzzy-logic models is very simple and there is no need to define the complex physicochemical reactions and time consuming mathematical formulations to predict the BOD value. This model, not only provides higher accuracy but also helps in fast monitoring of effluents, which in turn enables to implement appropriate control and measures.

REFERENCES

[1].Delzer, G.C., McKenzie, S.W., 1999. Fıve-Day Bıochemıcal Oxygen Demand-1, U.S. | Geological Survey TWRIBook, vol. 9 ISBN 978-0-607-92182-3 [chapter A7.2 5/99]. | [2].Dogan E, Sengorur B, Koklu R (2009). Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique. J Environ Manag 90:1229–1235 | [3]. Honggui H, Junfei Q (2009) Biological oxygen demand (BOD) soft measuring based on dynamic neural network (DNN): A simulation study. In: proceedings of the 7th Asian control conference, Hong Kong, China, August 27–29 | [4]. S.Areerachakul and S.Sanguansintukul (2009) "A Comparison between the Multiple Linear Regression Model and Neural Networks for Biochemical Oxygen Demand Estimations", The Eight International Symposium on Natural Language Processing, Thailand | [5]Kaan Yetilmezsoy, et.al.2011, Artificial intelligence-based prediction models for environmental engineering, Neural Network World 3/11, 193-218 | [6] S.R. Jang, (1993) ANFIS Adaptive-network-based fuzzy inference systems, IEEE Transactions on Systems Man and Cybernetics 23 665–685. | [7] T. Takagi, M. Sugeno, (1985) Fuzzy identification of systems and its applications to modeling and control, IEEE Transactions on Systems Man and Cybernetics 15 116–132. | [8] Mingzhi Huang, et.al, (2012). Modeling a Paper-Making Wastewater Treatment Process by Means of an Adaptive Network-Based Fuzzy Inference System and Principal Component Analysis. Ind. Eng. Chem. Res., 51, 6166–6174 | [9]Ali A. Najah, et.al. (2010). Water quality prediction model utilizing integrated wavelet-ANFIS model with cross-validation Neural Comput & Applic 21(5):833–841 | [10].M.Z. Huang, (2009), Control rules of aeration in a submerged biofilm wastewater treatment process using fuzzy neural networks, Expert Systems with Applications 36 10428–10437. | [11].M. Bongards, (2001) Improving the efficiency of a wastewater treatment plant by fuzzy control and neural networks, Water Science and Technology 3 189–196. | [12] Hanife Sariet.al., (2012). Fuzzy-logic modeling of Fenton's strong chemical oxidation process treating three types of landfill leachates. Environ Sci Pollut Res, DOI 10.1007/s11356-012-1370-6 | [13] Mostafa Rezazadeh shirdar, et.al,. (2011).Application of ANFIS system in prediction of machining parameters, Journal of theoretical and applied information technology. 33 (1) [83-93] | [14].APHA 1992. Standard methods for the examination of water and waste water. American Public Health Association 18th Ed. Academic Press, Washington D.C. pp214- 218. | [15].Xin, Y. & Xiaogang, S. (2009) Linear regression analysis: theory and computing. Singapore: World Scientific. | [16] M. Buragohain and C. Mahanta (2008), "A novel approach for ANFIS modeling based on full factorial design" Applied Soft Computing 8(1); 609–625. | [17] Dastorani, M.T., A. Moghadamnia, J. Piri and M. Rico-Ramirez , 2009 "Application of ANN and ANFIS models for reconstructing missing flow data". Environ. Monitoring Assess. doi: 10.1007/s10661-009-1012-8 | [18] Kalyanaraman, S.B., S.Arulselvi and S.Akilandeswari, 2006.Anfis modeling for estimation of BOD values in tannery effluents.J.Curr.Sci. 9(1):387-392.